Plant Genetics and Genomics: Crops and Models 19

Andrew Doust
Xianmin Diao *Editors*

# Genetics and Genomics of Setaria

Springer

Plant Genetics and Genomics: Crops and Models

Volume 19

**Series Editor**
Richard A. Jorgensen

More information about this series at http://www.springer.com/series/7397

Andrew Doust • Xianmin Diao
Editors

# Genetics and Genomics of Setaria

Springer

*Editors*
Andrew Doust
Department of Plant Biology, Ecology, and Evolution
Oklahoma State University
Stillwater, OK, USA

Xianmin Diao
Institute of Crop Sciences
Chinese Academy of Agricultural Sciences
Haidian District, Beijing, China

ISSN 2363-9601 ISSN 2363-961X (electronic)
Plant Genetics and Genomics: Crops and Models
ISBN 978-3-319-45103-9 ISBN 978-3-319-45105-3 (eBook)
DOI 10.1007/978-3-319-45105-3

Library of Congress Control Number: 2016950031



Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

*Setaria* is a genus of panicoid grasses that utilizes the highly efficient $C_4$ photosynthetic pathway and is related to other $C_4$ grasses such as switchgrass, pearl millet, maize, and sorghum. The Setaria system comprises two species (considered subspecies by some authors), namely, wild green foxtail (*S. viridis*), one of the most widespread weeds on the planet, and its domesticated cousin, foxtail millet (*S. italica*), a drought-hardy and nutritious cereal important in China, India, and Africa. Together, the two species make up a remarkable system for investigating different aspects of plant biology, ranging from the processes of ecological differentiation, domestication, morphological and developmental change, genetic regulation, $C_4$ photosynthesis, breeding, and genome evolution. The aim of this book is to introduce the Setaria system to a wider audience, explore current research in Setaria, and provide protocols and guidance for crossing, mutant production, creation of genetic resources, transformation, and genetic analysis.

The wide latitudinal and ecological range of green foxtail and foxtail millet has led to population divergence and local adaptation to a variety of conditions. The model *S. viridis* accession, A10.1, is consistent in its growth form under controlled conditions but very sensitive to growth environment, making it ideal for examining the molecular basis of abiotic stress. Accession A10.1 is a small variety of green foxtail that can be grown in growth chamber, greenhouse, and field trials and needs no special growth conditions. In particular, the small physical size and rapid life cycle of A10.1, coupled with a small diploid genome, lend itself to genetic analyses such as those that have commonly been performed in *Arabidopsis*. Further model accessions are being developed that combine attributes of foxtail millet, especially non-seed shattering, with other desirable traits.

Multiple chapters in this volume speak to the utility of Setaria as a model system for $C_4$ grass biology, including $C_4$ photosynthesis, cell wall regulation, root and shoot regulation, root-microbe interactions, herbicide tolerance, and drought stress. These studies are enabled by genetic and genomic resources, including multiple genome sequences for foxtail millet and green foxtail, multiple sequenced diversity lines for population genetics, and genome-wide association studies, a renewed interest in creating mapping populations and mutant collections, and efficient

transformation techniques, including the promise of a spike dip protocol analogous to the floral dip protocol for transformation that revolutionized *Arabidopsis* genetic research. High-throughput sequencing (HTS) has been an important component in the development of several of these resources, including genome by sequencing and whole-genome sequencing of diversity lines, and rapid identification of candidate loci underlying quantitative trait loci (QTL) and mutant phenotypes. Coupled with HTS, new gene editing techniques are allowing rapid and efficient reverse genetic approaches, including testing candidate loci generated from GWAS, QTL mapping, and fine mapping in mutant populations.

An important aspect of the Setaria system is that it is part of a larger set of genetic model species in the grasses that allow inferences about gene and genome evolution that is simply not possible in any other family. These models include rice, maize, sorghum, and *Brachypodium* (*B. distachyon* and *B. stacei*), a pooid $C_3$ model grass much like Setaria in its small size and ease of use. In addition, there is the tetraploid genome of switchgrass (*Panicum virgatum*), the diploid genome of its close relative *Panicum hallii*, and multiple draft genomes in progress in other species. These genomes span the vast majority of grass diversity and allow unparalleled opportunities for investigating genome evolution and the genetic basis of morphological and physiological evolution. In addition, grass synteny allows basic research in these model systems to be quickly and efficiently translated into agronomically important crops such as maize, rice, and wheat.

Setaria is unique among these grass model systems because it encompasses both an important domesticated cereal and an emerging model system. The drought hardiness of Setaria makes it an attractive crop in parts of China, India, and Africa and an alternative to other cereals such as pearl millet and sorghum. The use of Setaria as a model system to understand drought stress will rapidly be translated into both better Setaria varieties and the possibility of better drought-hardy cereals in general. The small size and ease of genetic analysis in Setaria also make it the model of choice for understanding the genetics and physiology of $C_4$ photosynthesis, and Setaria is a key tool in the grand challenge of converting $C_3$ photosynthetic grasses like rice into highly efficient $C_4$ cereals to feed an ever-growing human population.

Finally, a note on nomenclature. Setaria is used in this volume to denote the Setaria system (both *Setaria italica* and its wild progenitor *Setaria viridis*), written without italicization and starting with a capital letter. The two species names are written in italics as is normal for Latin binomials. In addition, *S. viridis* has been referred to in past literature as both green millet and green foxtail, but we advocate the use of green foxtail as its correct common name, as it is not a millet cereal grain.

Stillwater, OK — Andrew Doust
Haidian, Beijing, P.R. China — Xianmin Diao

# Contents

# Contributors

**Beverly Jose Agtuca** Divisions of Plant Science and Biochemistry, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA

**Fernanda Plucani do Amaral** Divisions of Plant Science and Biochemistry, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA

**Jeffrey L. Bennetzen** Department of Genetics, University of Georgia, Athens, GA, USA

**Eduardo Blumwald** Department of Plant Sciences, University of California, Davis, CA, USA

**Christopher W. Brown** University of Newcastle, Newcastle, NSW, Australia

**Thomas P. Brutnell** Donald Danforth Plant Science Center, St Louis, MO, USA

**Caitlin S. Byrt** School of Agriculture, Food and Wine, Waite Research Institute, University of Adelaide, Urrbrae, SA, Australia

**Carla Coelho** Donald Danforth Plant Science Center, St Louis, MO, USA

**Henri Darmency** Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne Franche-Comté, Dijon, France

**Christophe Délye** Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne Franche-Comté, Dijon, France

**Katrien M. Devos** Institute of Plant Breeding, Genetics and Genomics (Dept. of Crop and Soil Sciences), and Dept. of Plant Biology, University of Georgia, Athens, GA, USA

**Xianmin Diao** Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, People's Republic of China

**José R. Dinneny** Department of Plant Biology, Carnegie Institution for Science, Stanford, CA, USA

**Andrew N. Doust** Department of Plant Biology, Ecology, and Evolution, Oklahoma State University, Stillwater, OK, USA

**Joyce Van Eck** Boyce Thompson Institute, Ithaca, NY, USA

**Maximillian Feldman** Donald Danforth Plant Science Center, St. Louis, MO, USA

**Kenji Fukunaga** Prefectural University of Hiroshima, Shobara, Japan

**Robert T. Furbank** Australian Research Council Centre of Excellence for Translational Photosynthesis, Plant Science Division, Research School of Biology, The Australian National University, Acton, ACT, Australia

**Christopher P. L. Grof** University of Newcastle, Newcastle, NSW, Australia

**John G. Hodge** Department of Plant Biology, Ecology, and Evolution, Oklahoma State University, Stillwater, OK, USA

**Pu Huang** Donald Danforth Plant Science Center, St Louis, MO, USA

**Guanqing Jia** Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, People's Republic of China

**Hui Jiang** Donald Danforth Plant Science Center, St. Louis, MO, USA

**Elizabeth A. Kellogg** Donald Danforth Plant Science Center, St. Louis, MO, USA

**Kazuhiro Kikuchi** Donald Danforth Plant Science Center, St. Louis, MO, USA

**Christopher J. Lambrides** School of Agriculture and Food Sciences, The University of Queensland, St Lucia, QLD, Australia

**Anthony P. Martin** University of Newcastle, Newcastle, NSW, Australia

**Kimberly Maxson-Stein** Donald Danforth Plant Science Center, St. Louis, MO, USA

**Mehanathan Muthamilarasan** Plant Molecular Genetics and Genomics, National Institute of Plant Genome Research (NIPGR), New Delhi, Delhi, India

**Duc Q. Nguyen** University of Newcastle, Newcastle, NSW, Australia

**William M. Palmer** University of Newcastle, Newcastle, NSW, Australia

**Minkyu Park** Department of Genetics, University of Georgia, Athens, GA, USA

**Kaitlin Pidgeon** Boyce Thompson Institute, Ithaca, NY, USA

**Manoj Prasad** Plant Molecular Genetics and Genomics, National Institute of Plant Genome Research (NIPGR), New Delhi, Delhi, India

**Peng Qi** Institute of Plant Breeding, Genetics and Genomics (Dept. of Crop and Soil Sciences), and Dept. of Plant Biology, University of Georgia, Athens, GA, USA

**Prasenjit Saha** Department of Plant Sciences, University of California, Davis, CA, USA

**Jose Sebastian** Department of Plant Biology, Carnegie Institution for Science, Stanford, CA, USA

**Christine Shyu** Donald Danforth Plant Science Center, St. Louis, MO, USA

**Gary Stacey** Divisions of Plant Science and Biochemistry, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA

**Kerry Swartwood** Boyce Thompson Institute, Ithaca, NY, USA

**Hao Wang** Department of Genetics, University of Georgia, Athens, GA, USA

**TianYu Wang** Institute of Crop Sciences, Chinese Academy of Agriculture Sciences, Beijing, China

**Xiaomei Wu** Institute of Plant Breeding, Genetics and Genomics (Dept. of Crop and Soil Sciences), and Dept. of Plant Biology, University of Georgia, Athens, GA, USA

Center for Plant Environmental Sensing, and College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, China

**Hongye Zhou** Department of Genetics, University of Georgia, Athens, GA, USA

# Editors Biographies

**Andrew Doust** is an Associate Professor of Plant Biology in the Department of Plant Biology, Ecology, and Evolution at Oklahoma State University. After obtaining his Ph.D. from the University of Melbourne, Australia, he served as a postdoctoral fellow both at the University of Missouri-St. Louis and the University of Florida. His general interests revolve around the genetic regulation of development and how that changes over evolutionary time, with a particular concentration on domestication in grasses.

**Xianmin Diao** is a Professor of Plant Genetics in the Institute of Crop Sciences at the Chinese Academy of Agricultural Studies. He holds a Ph.D. in plant physiology from the Institute of Botany at the Chinese Academy of Sciences and formerly served as a visiting scholar at the University of California, Berkeley. As the leading scientist for the foxtail and proso millet section of China's agricultural research system, he runs a lab focused on foxtail millet germplasm management, breeding, genetics, and genomics.

# Part I
# Evolution

# Chapter 1
# Evolution of *Setaria*

**Elizabeth A. Kellogg**

**Abstract** The genus *Setaria* includes almost 100 species of panicoid grasses. Within the subfamily Panicoideae it falls in the tribe Paniceae, subtribe Cenchrinae. Members of the subtribe are characterized by the presence of sterile branches in the inflorescence, often known as "bristles." Major clades of *Setaria* are geographically localized, with the African species falling in a distinct clade from the South American ones, which are in turn distinct from the Asian ones. Many species have become weedy and are distributed widely in warm areas throughout the world. Nearly all members of the genus share a chromosome base number of $x = 9$, similar to most other members of Paniceae, and polyploidy is common, with some species including tetraploid, hexaploid, and octoploid members. The crop species *S. italica* was domesticated from the weed *S. viridis*, and both share a genome designated as A. A second diploid weed, *S. adhaerens*, is genomically distinct, with a genome designated as B. The tetraploid *S. verticillata* includes both A and B genomes, while tetraploid *S. faberi* appears to be derived from two A-like ancestors.

**Keywords** *Setaria* • Paniceae • Panicoideae • Poaceae • Polyploidy • Bristle clade

## 1.1 Relationships of *Setaria* to Other Monocots

The genus *Setaria* is a member of subtribe Cenchrinae, tribe Paniceae, subfamily Panicoideae, family Poaceae, order Poales, in the commelinid clade of monocots (Soreng et al. 2015; Kellogg 2015). Because *Setaria* inherits the characters of each of these larger clades, we will consider each in turn, progressing from the most to the least inclusive.

E.A. Kellogg, Ph.D. (✉)
Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132, USA
e-mail: ekellogg@danforthcenter.org



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_1

### *1.1.1 Characteristics of Setaria Shared with the Commelinid Clade, Poales, and Poaceae*

#### 1.1.1.1 Commelinids

The commelinid clade includes Poales, Zingiberales (gingers and bananas), and Commelinales (spiderworts and their relatives). All members of this clade have endosperm that is well developed and persists in the seed (Stevens 2012). In addition, the commelinid monocots have unique cell walls. The hemicelluloses in the primary walls are largely arabinoxylans, specifically glucuronoarabinoxylans (Carpita 1996; Withers et al. 2012); the latter appear to be unique to the commelinids. The amount of pectin and protein is fairly low in comparison to eudicots (Carpita 1996). In the secondary walls, major lignin subunits are *p*-coumaric and ferulic acid (Harris and Hartley 1980; Harris and Tretheway 2009). Much of the information on the biochemistry of the commelinid secondary wall comes from grasses (Withers et al. 2012; Petrik et al. 2014; Molinari et al. 2013) but is presumed to apply to all commelinids. All commelinids accumulate silica in their leaves (Stevens 2012). Thus, *Setaria* could be a useful model for studies of endosperm, cell walls, and silica accumulation, with the results of such studies applying not only to grasses, but to other members of the commelinid clade.

#### 1.1.1.2 Poales

The Poales in its current broad sense includes 16 families (Angiosperm Phylogeny Group 2009; Givnish et al. 2010), all of which accumulate silica specifically in the leaf epidermis (Stevens 2012). Silica accumulation protects the plant from pathogenic bacteria and fungi (Isa et al. 2010; Ma and Yamaji 2006), and also appears to reduce insect herbivory (Massey and Hartley 2009; Garbuzov et al. 2011). In addition, deposition of silica provides structural support, reduces the uptake of toxic metals, and regulates water loss (Isa et al. 2010; Ma and Yamaji 2006). One popular theory suggested that production of silica was selected as a defense against mammalian grazers because it would wear down their teeth (Simpson 1951; Baker et al. 1959). However, this idea is not supported by data (Sanson et al. 2007; Strömberg 2006).

Endosperm development in all Poales is unique among angiosperms, with multiple rounds of nuclear division before cell walls form (nuclear endosperm) (Stevens 2012). Few other morphological characters are shared by members of Poales, although many are wind pollinated, often occur in nutrient-poor habitats, and are often fire adapted (Linder and Rudall 2005). Flowers generally occur in tiny clusters called “spikelets” in both Cyperaceae and Poaceae, but the structure of these is quite different and nonhomologous in the two families.

Within Poales, Poaceae fall into the graminid clade, a well-supported group that also includes Flagellariaceae, Restionaceae (which now includes members of the former Centrolepidaceae), Anarthriaceae, Joinvilleaceae, and Ecdeiocoleaceae.

Members of the graminid clade have monoporate pollen, with a raised ring or annulus around the pore (Stevens 2012), a character that is retained in *Setaria*. The functional significance of this pollen form is unknown. Nearly all graminids have two-ranked (distichous) sheathing leaves. The endothecium of the anther has girdle-like thickenings. Stigmas are generally plumose, with receptive cells on multicellular branches. The graminids also all share the ability to produce flavones.

The immediate sister group of Poaceae is uncertain. Possible candidates are *Joinvillea*, the sole genus in Joinvilleaceae, or Ecdeiocoleaceae, a family with two genera, *Ecdeiocolea* and *Georgeantha*. Current data suggest that the two families are sisters, and that the clade is then sister to Poaceae (McKain et al. 2016). Members of both families have conventional monocot flowers with two whorls of perianth parts, and thus their structure sheds little light on the homologies of the grass spikelet (but see (Preston et al. 2009; Kellogg 2015)). Like *Setaria* and other grasses, both Joinvilleaceae and Ecdeiocoleaceae have dumbbell-shaped stomatal guard cells. This guard cell shape is thought to enhance the speed of pore opening (Haworth et al. 2011; Franks and Farquhar 2007). Also Joinvilleaceae shares with the grasses the pattern of alternating long and short cells in the epidermis (Campbell and Kellogg 1987). As with many such morphological characteristics, the genetic controls and functional significance of this character are unknown.

#### 1.1.1.3 Poaceae

Poaceae, or Gramineae (both names are correct), is the most speciose of the families in Poales. It includes ca. 12,000 species (Clayton et al. 2006 onwards; Kellogg 2015) and is clearly monophyletic (Kellogg and Campbell 1987; Kellogg and Linder 1995; Vicentini et al. 2008; GPWG 2001; GPWG II 2012).

The grasses all share a distinctive embryo and fruit (GPWG 2001; Kellogg 2015). The seed coat is generally fused to the inner epidermis of the pericarp, forming a single seeded fruit or caryopsis (Fig. 1.1a). Unlike all other commelinid monocots (and indeed most monocots), the grass embryo is highly differentiated (Kellogg 2000; Campbell and Kellogg 1987; Rudall et al. 2005), with a well-developed shoot apical meristem surrounded by a sheath-like structure, the coleoptile, and bearing two or more leaves (Sylvester et al. 2001) (Fig. 1.1a). The root apical meristem is also differentiated and surrounded by a coleorhiza. Attached to the embryo is a large shield-shaped haustorial organ, the scutellum. Together the coleoptile and scutellum appear to represent the sheath and blade, respectively, of a highly modified cotyledon (Takacs et al. 2012).

Also characteristic of Poaceae is the formation of tiny trichomes (microhairs) from the short cells of the leaf epidermis. Microhairs are two-celled, with an elongate apical cell (Johnston and Watson 1976). The functional significance of microhairs is unknown, although it is possible that they could be secretory in some instances. While the ability to produce microhairs is clearly ancestral in the grasses (GPWG 2001) and occurs in *Setaria* as well as all other panicoid species, the cool-season grasses in subfamily Pooideae do not produce them so they are not universal in grasses.

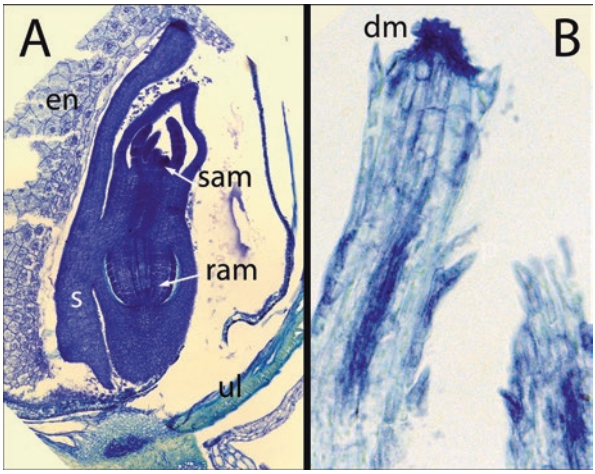


**Fig. 1.1** (**a**) Embryo of *Setaria viridis*. As in all other grasses, the *Setaria* embryo has well-developed shoot and root meristems and a clear scutellum. The deep cleft between the base of the scutellum and the coleorhiza is common in most grasses except for the Pooideae. en, endosperm; sam, shoot apical meristem; ram, root apical meristem; s, scutellum. (**b**) Apex of young bristle showing collapsed cells. Bristles appear to lose their meristems early in development. dm, degenerating meristem. Photos by John G. Hodge

Poaceae genomes have been studied extensively because genomic information is so essential for breeding efforts. A whole genome duplication occurred in the common ancestor of all grasses, so that many loci are retained in duplicate (Goff et al. 2002; Yu et al. 2002; Paterson et al. 2004, 2009; McKain et al. 2016).

Poaceae is divided into 12 subfamilies (Fig. 1.2) (Kellogg 2015; Soreng et al. 2015). The ones that diverged early in the evolution of the family include only a handful of species (Anomochlooideae, four species, Pharoideae, 12, and Puelioideae, 11) (Kellogg 2015). The vast majority of species fall in the BOP and PACMAD clades, the names of which are acronyms for the included subfamilies. BOP includes Bambusoideae, Oryzoideae, and Pooideae, whereas PACMAD includes Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae.

The grass spikelet, which is a tiny spike delimited by two bracts (glumes) and with one or more flowers, characterizes all grasses except Anomochlooideae. The precise timing of origin of the spikelet, however, is unclear. Either the spikelet was present in the common ancestor of all grasses and was then highly modified in Anomochlooideae, or the spikelet originated in the common ancestor of Pharoideae and all remaining grasses (GPWG 2001; Preston et al. 2009; Kellogg 2015) (node 1, Fig. 1.2). In either case, *Setaria* is like all but about four species of grasses in having flowers borne in spikelets.

Other widespread aspects of grasses characterize Puelioideae plus the BOP+PACMAD clades (i.e., descendants of node 2, Fig. 1.2), and not

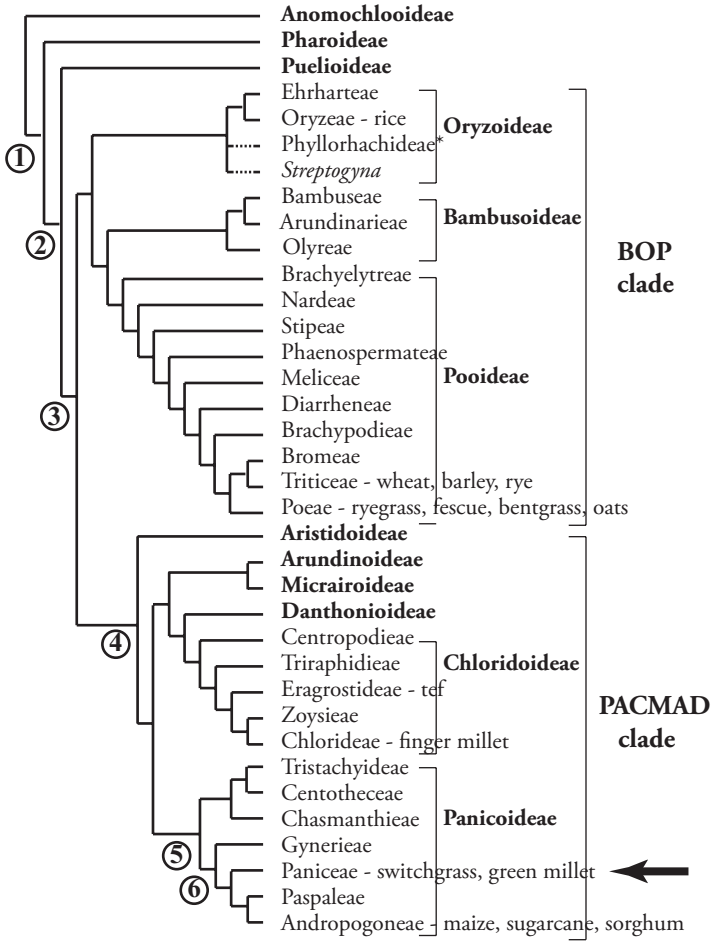


**Fig. 1.2** Phylogeny of the grasses, based largely on GPWG II (2012) and redrawn from Kellogg (2015). Arrow points to subfamily Panicoideae, tribe Paniceae, which includes *Setaria*

Anomochlooideae or Pharoideae. For example, style branches and stigmas are reduced to two in this group (although the character reverses in some taxa), and the stigmas have two orders of branching (GPWG 2001). Spikelets each have multiple flowers (another character that reverses frequently). Anther walls have a middle layer that breaks down during development, and the inner walls of the endothecial cells become fibrous at maturity.

The female gametophyte in the grasses is fairly conventional in early development, with an egg, two synergids, a binucleate central cell, and antipodal cells. However, in all investigated species other than the early diverging genera *Streptochaeta* and *Pharus* (Sajo et al. 2007, 2008), the antipodals continue to divide (Anton and Cocucci 1984; Evans and Grossniklaus 2009; Shadowsky 1926). The function of these extra divisions is unknown.

The BOP+PACMAD clade (descendants of node 3, Fig. 1.2) has no obvious morphological synapomorphy. The Grass Phylogeny Working Group (GPWG 2001) suggested that lack of a pseudopetiole in the leaves, reduction of lodicule number to 2 and stamen number to 3 might be synapomorphic. Although these characters reverse in a number of lineages, they all characterize *Setaria*.

Members of the PACMAD clade (node 4, Fig. 1.2) have nothing obvious in common. They are thought to share an elongated mesocotyl internode in the embryo, but relatively few species have actually been investigated for this character and it is unclear how reliable or consistent it is (GPWG 2001). The clade also includes all 24 origins of the $C_4$ photosynthetic pathway in grasses (GPWG II 2012).

## *1.1.2 Characteristics of Setaria Shared with Subfamily Panicoideae, Tribe Paniceae, and Subtribe Cenchrinae*

### 1.1.2.1 Panicoideae

Nearly 1/3 of the species of Poaceae are in subfamily Panicoideae. Panicoideae s.s. (node 6, Fig. 1.2) was one of the earliest subfamilies to be recognized as distinct. In 1810, Robert Brown noted that the group (which he called Paniceae) mostly has spikelets with exactly two flowers, with the upper one bisexual and the lower one staminate or sterile (Brown 1810, 1814). Recent phylogenetic work has shown that Panicoideae s.s. is part of a larger clade which now bears the name Panicoideae (node 5, Fig. 1.2), in which the spikelet morphology is more variable (Sánchez-Ken and Clark 2007, 2010).

Spikelets in Panicoideae s.s. are dorsiventrally compressed. The glumes and lemmas are generally not folded and are borne ab- and adaxially in relation to the spikelet-bearing axis. This pattern of compression contrasts with that of most other grasses such as rice, tef, and *Brachypodium*, in which the glumes and lemma are both folded along the midrib, a pattern known as lateral compression. In these taxa, the glumes and lemmas initiate at right angles to the spikelet-bearing axis. As with many such morphological characters, the significance of this highly consistent difference is unknown.

Spikelet development in the panicoids is basipetal, with the distal flower maturing before the proximal one (Bess et al. 2005; Doust and Kellogg 2002; Malcomber and Kellogg 2004). This pattern is similar to that in rice, but distinct from what is found in Pooideae, Chloridoideae, and other major groups.

Silica bodies in the leaf epidermis of panicoid grasses are generally bilobed in surface view and symmetrical in cross section (Piperno 2006; Piperno and Pearsall 1998). Nothing is known about deposition of silica in the grass epidermis and the mechanism by which silica body shape is defined.

Early phylogenetic work in subfamily Panicoideae found that the phylogeny reflected chromosome numbers rather than photosynthetic pathway as had been thought

previously (Gómez-Martinez and Culham 2000; Aliscioni et al. 2003; Giussani et al. 2001). The ancestral base chromosome number of the subfamily is unknown but most likely to be 11 or 12, and the number was then reduced in the common ancestor of Panicoideae s.s. One descendant of this ancestor acquired a base number of $x=9$, a number that now characterizes the tribe Paniceae, whereas the other descendant acquired a number of $x=10$, which is shared by the tribes Andropogoneae and Paspaleae.

#### 1.1.2.2 Paniceae

Within Paniceae, major clades are strongly supported by both nuclear, chloroplast, and mitochondrial sequences (Vicentini et al. 2008; GPWG II 2012; Washburn et al. 2015). All analyses to date have identified clades corresponding to subtribes Cenchrinae, Melinidinae, Panicinae, Boivinellinae, Neurachninae, and Anthephorinae (Fig. 1.3). The genus *Dichanthelium* is monophyletic and could be placed in its own subtribe (Dichantheliinae (Soreng et al. 2015)). In addition, there is an unnamed clade made up of the genera *Sacciolepis*, *Trichanthecium*, and *Kellochoa* plus a number of species formerly placed in *Panicum* (node 4, Fig. 1.3) (Morrone et al. 2012; GPWG II 2012; Zuloaga et al. 2011; Nicola et al. 2015).

Cenchrinae, Melinidinae, and Panicinae form a robust group (the MPC clade) (Morrone et al. 2012; GPWG II 2012) (node 2, Fig. 1.3), with Cenchrinae and Melinidinae sisters (node 3, Fig. 1.3 (Washburn et al. 2015)). The clade was first identified as the "$C_4$ three subtypes" clade by Giussani et al. (2001) because all members are $C_4$, but each subtribe exhibits a different subtype of the $C_4$ pathway. Cenchrinae includes species that are NADP-ME subtype, Melinidinae members are PCK, and Panicinae are NAD-ME. Each $C_4$ subtype has characteristic leaf anatomy (Hattersley 1987; Hattersley and Watson 1992; Prendergast and Hattersley 1987; Prendergast et al. 1987). In Panicinae and Melinidinae, each vein is surrounded by an inner sheath of thick walled cells, the mestome sheath, and an outer sheath of parenchymatous cells. Carbon reduc-

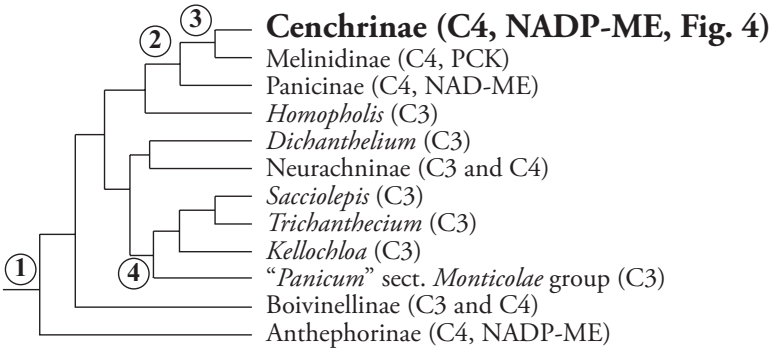


**Fig. 1.3** Phylogeny of Paniceae based on chloroplast, mitochondrial, and nuclear rDNA gene sequences (Washburn et al. 2015; Nicola et al. 2015). The clade made up of Cenchrinae, Melinidinae, and Panicinae is found in all molecular phylogenies. Other relationships, particularly those surrounding node 1, are contradicted by other gene trees

tion occurs in the latter. In Cenchrinae, and thus in *Setaria*, as in most NADP-ME grasses, veins are surrounded by a single sheath, a derived condition in the subtribe.

Relationships among the other clades of Paniceae are unclear. Chloroplast data place *Dichanthelium* as sister to the MPC clade, whereas nuclear data place it sister to all other Paniceae (Vicentini et al. 2008) and combined chloroplast, mitochondrial, and nuclear RNA data place it sister to Neurachninae (Washburn et al. 2015). Conversely, Anthephorinae is placed sister to all Paniceae by chloroplast data (node 1, Fig. 1.3), but sister to the MPC clade by nuclear genes.

#### 1.1.2.3 Cenchrinae

The Cenchrinae is also known as the "bristle clade," because almost all members of the clade form sterile branches ("bristles") in the inflorescence. These sterile branches originate as ordinary branches, but instead of forming spikelets, they grow out and terminate blindly (Doust and Kellogg 2002). The bristles may form a meristem at their apex, but this often simply aborts, leaving a small collapsed set of cells (Fig. 1.1b). Bristles may be restricted to the ends of branches, or a bristle may be paired with each spikelet, or individual spikelets may be surrounded by an involucre of bristles. In the latter case, the bristles may be terete or flattened.

Two species, *Zuloagaea bulbosa* and "*Panicum*" *antidotale*, lack bristles. (The latter species is unrelated to true *Panicum*, which is in Panicinae, but has not yet been transferred to another genus.) Developmental studies in *Zuloagaea* show that early development in that species is strikingly similar to that of *Panicum miliaceum* (*Panicum* s.s.) and there is no evidence of bristle formation at any point in development (Bess et al. 2005, 2006).

The phylogeny of the group is poorly resolved largely because no one has yet investigated it using a sufficient number of markers. Nonetheless, a few strong clades can be identified. The *Cenchrus* clade (*Cenchrus* sensu lato, Fig. 1.4) includes both *Cenchrus* and the former genus *Pennisetum*, plus the monotypic *Odontelytrum* (Chemisquy et al. 2010; Donadio et al. 2009; Morrone et al. 2012). All species of *Cenchrus* s.l. form an abscission zone at the base of the primary branch such that the spikelets fall from the plant surrounded by an involucre of bristles. Developmentally, the species are also distinct because the primary branch enlarges isodiametrically, rather than growing primarily in a proximo-distal direction (Doust & Kellogg 2002). While many species of *Cenchrus* s.s. are easily identified by their flattened bristles forming an involucre around the spikelets, others intergrade morphologically with the former *Pennisetum*. Thus, the boundary between the genera is not sharp, consistent with the pattern found in phylogenetic studies (Donadio et al. 2009; Chemisquy et al. 2010).

Primary branches of the inflorescence are spirally arranged in most species, whereas the secondaries and higher order branches are distichous (Bess et al. 2005; Doust and Kellogg 2002; Kellogg et al. 2004, 2013).

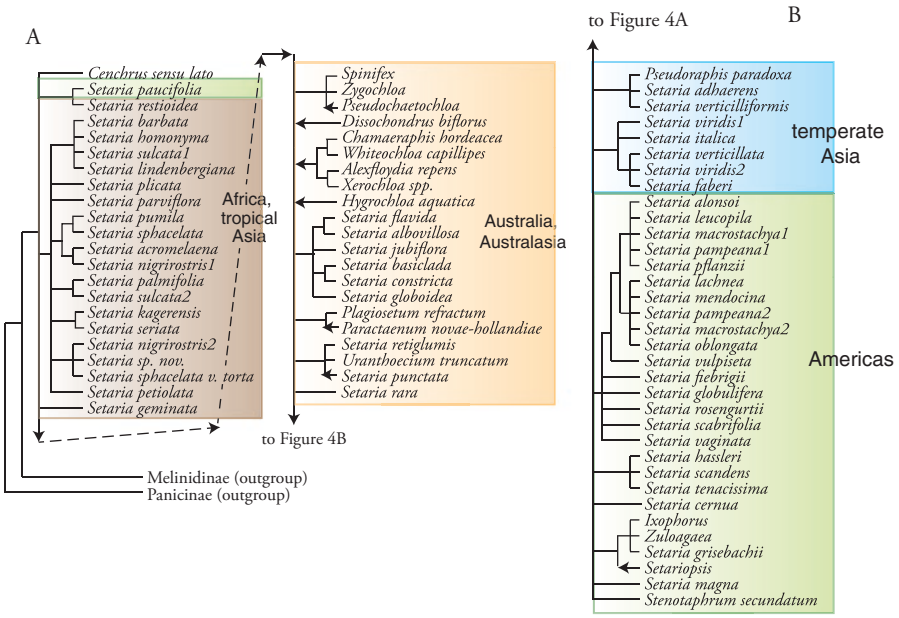


**Fig. 1.4** Phylogeny of Cenchrinae based on the chloroplast gene *ndhF*, focusing on species in *Setaria*. Redrawn from Kellogg et al. (2009). *Arrows* show the approximate placement of additional taxa included in the study of GPWG II (2012). All branches shown have either parsimony or maximum likelihood bootstrap values >80, or Bayesian posterior probability >0.95, or both. *Brown* box indicates taxa native to Africa or tropical Asia, *yellow* is Australia and Australasia, *blue* is temperate Asia and *green* is the Americas, mostly South and Central America

## 1.2 Relationships Within the Genus *Setaria*

### *1.2.1 Phylogeny and Characteristics of the Genus*

Species of *Setaria* are described in detail in two monographs, which together cover 99 species (Morrone et al. 2014; Pensiero 1999), although Clayton et al. (2006 onward) lists 103 names. The genus has no unique character and as currently defined is likely to be para- or polyphyletic. The current phylogeny shows a number of well-supported clades corresponding largely to geography, but relationships among them are unresolved, making generic circumscription impossible at the moment (Fig. 1.4). As currently circumscribed, *Setaria* includes the members of Cenchrinae that do not fall in the *Cenchrus* clade, have bisexual spikelets (i.e., are not *Spinifex* or *Zygochloa*), and lack the distinctive characters of the various oligotypic genera such as *Dissochondrus*, *Paractaenum*, or *Plagiosetum*. Almost certainly, *Setaria* will need to be expanded to include some of these elements, but without a solid phylogeny it is hard to find a good rationale for doing so.

Species of *Setaria* occur in warm regions throughout the world, and in diverse habitats (Morrone et al. 2014). Some species, such as *S. sulcata* and *S. palmifolia*, occur in

disturbed areas in moist forest shade. Others, such as *S. nigrirostris* and *S. sphacelata*, are found in damp grasslands, and still others, such as *S. rara* and *S. reflexa*, in dry open habitats. A handful of species, notably *S. viridis* and *S. pumila*, are weedy and have followed human activity to spread far beyond their original distribution.

The number and position of bristles in the inflorescence varies considerably among *Setaria* species (Morrone et al. 2014). In species such as *Setaria palmifolia*, each spikelet is accompanied by a single bristle. In other species, such as *S. parviflora* and *S. viridis*, each mature spikelet is surrounded by multiple bristles. The relationship between the number of spikelets and the number of bristles is developmentally complex however (Doust and Kellogg 2002). Bristle and spikelet identity are specified early in inflorescence development. In some cases, all spikelets develop to maturity so that the number of bristles per spikelet reflects meristem identity decisions. In other species, however, late forming spikelets fail to develop so that high numbers of bristles per spikelet reflect a process of spikelet abortion rather than branch identity specification.

In some species of *Setaria* the primary branches of the inflorescence are themselves unbranched (i.e., the spikelets are borne directly on the primary branches) and the branches end in a sharp bristle-like tip (Morrone et al. 2014). The inflorescence thus looks superficially similar to that in *Paspalum*, but the presence of the terminal bristle is diagnostic; species with this inflorescence morphology have often been placed in a genus *Paspalidium*. However, species occur in which some spikelets are associated with bristles in addition to the one at the branch tip and thus the morphology intergrades with that of *Setaria* sensu stricto. Recognizing this morphological intermediacy, all *Paspalidium* species have been transferred to *Setaria* (Veldkamp 1994; Webster 1993, 1995), and the transfer has been supported by phylogenetic data (Morrone et al. 2012; Kellogg et al. 2009; GPWG II 2012).

The species of *Setaria* vary widely in inflorescence architecture and leaf form (Morrone et al. 2014). Inflorescences may be narrow with short stiff lateral branches (the inflorescence thus shaped like a bottle brush, e.g., *S. pumila*, *S. sphacelata*, *S. nigrirostris*), broad and lax with spreading branches (shaped like a Christmas tree, e.g., *S. grandis*, *S. sulcata*, *S. lindenbergiana*), or sparse with few spreading primary branches (like an antenna, e.g., *S. jubiflora*, *S. flavida*). Each primary branch may produce spikelets directly (e.g., *S. jubiflora*, *S. flavida*, *S. rara*) or may rebranch up to six times (e.g., *S. parviflora*, *S. pumila*). Plants may be annual (e.g., *S. faberi*, *S. acromelaena*, *S. sagittifolia*, *S. viridis*) or perennial (most species), and may be a few cm (e.g., some specimens of *S. clementii*, *S. ustilata*) to over 1 m (e.g., *S. grandis*) tall. Spikelets are generally ovate but sometimes may be elongate or orbicular. Leaves are generally flat, but some species (e.g., *S. sulcata*, *S. palmifolia*) have striking folded leaves. The latter were once placed in their own section because the leaf morphology is so distinctive, but they do not form a clade in molecular phylogenies. Sagittate leaves are found in *S. sagittifolia* and *S. appendiculata*.

As in all groups of grasses, polyploids are common (Table 1.1). Except for polyploids involving *Setaria viridis* (see next section) the history of few of these has been disentangled, although it would be straightforward to do so using low-copy nuclear genes. Sequences of the nuclear gene *Knotted1* have shown that *S. flavida*

**Table 1.1** Published chromosome numbers for species of *Setaria*

| Species | Origin | *n* | Reference | 2*n* | Reference |
|---|---|---|---|---|---|
| *adhaerens* | Asia | 9 | Gupta and Singh (1977) | | |
| *apiculata* | Australia | | | 36 | Le Thierry d'Ennequin et al. (1998)[a] |
| *barbata* | Africa | 18 | Olorode (1975) | 54 | Gadella (1977) |
| | | 27 | Christopher and Abraham (1976) and Dujardin (1978) | 56 | Sarkar et al. (1976) |
| | | 28 | Sarkar et al. (1976) | | |
| *faberi* | Asia | | | 36 | Probatova and Sokolovskaya (1983) and Warwick et al. (1987, 1997) |
| *fiebrigii* | South America | 18 | Oliveira Freitas-Sacchet (1980) and Oliveira Freitas-Sacchet et al. (1984) | 36 | Pensiero (1999) |
| *flavida* | Australia | 18 | Bir and Chauhan (1990) | 44 | Sharma and Sharma (1979) |
| | | 27 | Mehra (1982), Bir and Sahni (1983) and Nadeem Ahsan et al. (1994) | 54 | Sinha et al. (1990) |
| | | 56 | Bir and Chauhan (1990) | | |
| *geminata* | Africa | 9 | Rao and Mwasumbi (1981) and Nadeem Ahsan et al. (1994) | | |
| *grisebachii* | Central and South America | | | 18 | Reeder (1971) |
| *homonyma* | Africa | 10 | Singh and Gupta (1977) | | |
| | | 18 | Mehra and Sharma (1975) | | |

(continued)

**Table 1.1** (continued)

| Species | Origin | *n* | Reference | 2*n* | Reference |
|---|---|---|---|---|---|
| *italica* | Asia | 9 | Khosla and Sharma (1973), Gupta and Singh (1977), Mehra (1982) and Sinha et al. (1990) | 18 | Christopher and Abraham (1976), Li and Chen (1985) and Sinha et al. (1990), Chikara and Gupta (1979), Frey et al. (1981), Zhou et al. (1989), Kozuharov and Petrova (1991), Li et al. (1996), Wu and Bai (2000), and Le Thierry d'Ennequin et al. (1998)[a] |
| | | | | 36 | Li and Chen (1985) |
| *kagerensis* | Africa | | | 18 | Lakshmi and Yacob (1978) |
| *lachnea* | South America | 18 | Gupta and Singh (1977) | 36 | Bowden and Seen (1962), Manero de Zamelzú and Ochoa de Suárez (1991), Pensiero (1999), and Le Thierry d'Ennequin et al. (1998)[a] |
| *leucopila* | SW US, Mexico, South America | | | 54, 68, 72 | Emery (1957a) |
| *longiseta* | Africa | 18 | (Olorode 1975) | 36 | |
| *macrostachya* | North, Central, and South America | 27 | Gupta and Singh (1977) | 54 | Emery (1957b), Pensiero (1999), and Le Thierry d'Ennequin et al. (1998)[a] |
| | | | | 72 | Gupta and Singh (1977) |
| *magna* | North, Central, and South America | | | 36 | Brown (1948) |
| *nigrirostris* | Africa | 9 | Gupta and Singh (1977) | 18 | Raman et al. (1959) and Le Thierry d'Ennequin et al. (1998)[a] |
| | | 18 | Spies and duPlessis (1986) | 36, 54 | Spies and duPlessis (1986), Raman et al. (1959), and Le Thierry d'Ennequin et al. (1998)[a] |
| | | 27 | Spies and duPlessis (1986) | | |

(continued)

**Table 1.1** (continued)

| Species | Origin | *n* | Reference | 2*n* | Reference |
|---|---|---|---|---|---|
| *oblongata* | Argentina, Bolivia | 18 | Tiranti and Genghini (2000) | | |
| *palmifolia* | Asia, Africa | 27 | Mehra and Sharma (1975), Mehra (1982) and Christopher and Abraham (1976) | 54 | Christopher and Abraham (1976) |
| | | | | 36 | Le Thierry d'Ennequin et al. (1998)[a] |
| *pampeana* | Argentina | | | ca. 50 | Pensiero (1999) |
| *parviflora* | North, Central, and South America | 18 | Gupta and Singh (1977), Oliveira Freitas-Sacchet (1980) and Mehra (1982) | 36 | Gould and Soderstrom (1967), Pohl and Davidse (1971), Norrmann et al. (1994), and Le Thierry d'Ennequin et al. (1998)[a] |
| | | | | 72 | Gould and Soderstrom (1967) and Fernández and Queiróz (1969) |
| *pflanzii* | South America | | | 36 | Caponio and Pensiero (2002) |
| *plicata* | Asia | 36 | Mehra (1982) | | |
| *pumila* | Africa, Asia | 9, 18+0–2B, 27 | Sahni (1989) | 36 | Sahni (1989), Kozuharov and Petrova (1991), Baltisberger (1988), Devesa et al. (1991), and Singh and Godward (1960) |
| | | 36 | Sahni (1989) and Nadeem Ahsan et al. (1994) | 54 | Le Thierry d'Ennequin et al. (1998)[a] |
| *restioidea* | Africa | 18 | Dujardin (1978) | | |
| *rosengurtii* | South America | 54 | Oliveira Freitas-Sacchet et al. (1984) and Oliveira Freitas-Sacchet (1980) | | |

(continued)

**Table 1.1** (continued)

| Species | Origin | *n* | Reference | 2*n* | Reference |
|---|---|---|---|---|---|
| *scabrifolia* | South America | 18 | Oliveira Freitas-Sacchet et al. (1984) and Oliveira Freitas-Sacchet (1980) | | |
| *sphacelata* | Africa | 9 | Gupta and Singh (1977), Dujardin (1978) and Rao and Mwasumbi (1981) | 18 | deWet (1958) |
| | | 18 | Gupta and Singh (1977), Bir and Sahni (1986, 1987) and Sahni (1989) | 36 | deWet (1954), deWet and Anderson (1956), and Le Thierry d'Ennequin et al. (1998)[a] |
| | | 27 | Gupta and Singh (1977) | 54 | Gupta and Singh (1977) and deWet (1958) |
| *sulcata* | Central and South America | | | 18 | Gupta and Singh (1977) |
| | | 16 | Oliveira Freitas-Sacchet et al. (1984) and Oliveira Freitas-Sacchet (1980) | 32 | Oliveira Freitas-Sacchet (1980) |
| | | 18 | Olorode (1975) and Dujardin (1978) | 36 | Quarín (1977), deWet and Anderson (1956), and Le Thierry d'Ennequin et al. (1998)[a] |
| | | | | 54 | Moffett and Hurcombe (1949) |
| *tenacissima* | Central and South America | | | 36 | Sede et al. (2010) |
| *vaginata* | South America | 18 | Oliveira Freitas-Sacchet et al. (1984) and Oliveira Freitas-Sacchet (1980) | | |

(continued)

**Table 1.1** (continued)

| Species | Origin | *n* | Reference | 2*n* | Reference |
|---|---|---|---|---|---|
| *verticillata* | Eurasia | 9 | Christopher and Abraham (1976)[b] | 18 | deWet (1954) and Wu and Bai (2000)[b] |
| | | 18 | Gupta and Singh (1977), Sahni (1989), Bir and Sahni (1986) and Bala and Sachdeva (1989, 1990) | 36 | Váchová and Feráková (1980) |
| | | 27 | Christopher and Abraham (1976), Gupta and Singh (1977), Mehra (1982), Sahni (1989), Bir and Sahni (1983), Sahni and Bir (1985), Bala and Sachdeva (1989, 1990) and Sinha et al. (1990) | 54 | Khosla and Sharma (1973), Gupta and Singh (1977), and Sinha et al. (1990) |
| | | 36, 54 | Sahni (1989) and Bir and Sahni (1986) | | |
| *viridis* | Asia | 9 | Christopher and Abraham (1976), Gupta and Singh (1977) and Koul and Gohil (1988, 1991) | 18 | Khosla and Sharma (1973), Chopanov and Yurtsev (1976), Magulaev (1976), Váchová (1978), Kliphuis and Wieffering (1979), Belaeva and Siplivinsky (1981), Löve and Löve (1981), Guzik (1984), Li and Chen (1985), Kozuharov and Petrova (1991), Xu et al. (1992), Moffett and Hurcombe (1949), Le Thierry d'Ennequin et al. (1998),[a] and (Layton and Kellogg 2014)[a] |
| | | 18 | Löve and Löve (1981), Saxena and Gupta (1969) and Mulligan (1984) | | |

(continued)

**Table 1.1** (continued)

| Species | Origin | *n* | Reference | 2*n* | Reference |
|---|---|---|---|---|---|
| *vulpiseta* | Central and South America | 9 | Sede et al. (2010) | 36 | Pensiero (1999) |
| | | 18 | Oliveira Freitas-Sacchet et al. (1984) and Oliveira Freitas-Sacchet (1980) | 54 | Norrmann et al. (1994) |

Table reproduced and updated from Kellogg et al. (2009). Names of species follow Morrone et al. (2014) and Pensiero (1999) and have been updated from those in the original publications
[a]Estimated by flow cytometry
[b]Likely misidentified specimens of *S. adhaerens*

and *S. jubiflora* are the products of a single polyploidization event (Doust et al. 2007). One of the genomes that produced the tetraploid ancestor of the two species also appears to be shared with *S. grisebachii* (diploid) and also with *Stenotaphrum secundatum* (a presumed diploid), *Ixophorus unisetus* (tetraploid), and *Zuloagaea bulbosa* (tetraploid). The origin of the other genome is less clear.

A handful of published chromosome counts appear to have a base number other than $x=9$ (Table 1.1). One accession of *S. homonyma* is reported to have $n=10$ (Singh and Gupta 1977), one accession of *S. sphacelata* is apparently $n=18+1B$ (Dujardin 1978); chromosome spreads are illustrated in both papers, and the counts appear accurate, although Singh and Gupta acknowledge that the *S. homonyma* count might be 9+1B. An accession of *S. sulcata* may have $n=16$ and $2n=32$ (Oliveira Freitas-Sacchet et al. 1984; Oliveira Freitas-Sacchet 1980), but the count is not documented photographically.

### *1.2.2 The Temperate Asian Clade*

The annual species *Setaria italica*, *S. viridis*, *S. faberi*, and *S. verticillata* form a clade in chloroplast phylogenies (Kellogg et al. 2009; Layton and Kellogg 2014) and in phylogenies using the nuclear genes encoding *knotted1* and 5S rDNA (Zhao et al. 2013; Layton and Kellogg 2014; Doust et al. 2007) (Fig. 1.4). *S. viridis* is the wild ancestor of the cultivated species *S. italica*, as documented by data from many sources, and the two remain interfertile (Le Thierry d'Ennequin et al. 2000; Hunt et al. 2008; Hirano et al. 2011; Darmency et al. 1987; Shi et al. 2008; Huang et al. 2014). This relationship is discussed more extensively in the chapters by Jia (Chap. 2), Huang and Feldman (Chap. 3), and Diao and Jia (Chap. 4) in this book.

The diploid genome of *S. italica* was designated as the A genome by Li et al. (1945); diploid *S. viridis* shares the A genome with *S. italica*, as verified by hybrid fertility, and cytogenomic, enzymatic, and molecular markers (Li et al. 1945; Wang

et al. 1998; Benabdelmouna et al. 2001a, b; Darmency and Pernes 1987). The diploid genome of *S. adhaerens* is distinct from that of *S. viridis* and *S. italica* and has been designated as the B genome by Benabdelmouna et al. (2001b); this designation has been confirmed by Wang et al. (2009) and Zhao et al. (2013). In addition, sequence data show that *S. adhaerens* and *S. viridis* are not closely related (Fig. 1.4) (Layton and Kellogg 2014). The diploid genome of *S. grisebachii* from America was identified as genome C due to its poor hybridization signals with both the A genome of *S. viridis* and the B genome of *S. adhaerens* (Wang et al. 2009).

A combination of molecular phylogenetics and cytogenetics has identified the progenitors of several polyploid taxa. Chromosomes of the tetraploid species *S. pumila* (often erroneously known as *S. glauca* (Morrone et al. 2014)) and *S. parviflora* strongly cross-hybridized but no hybridization signal was detected when the chromosomes of these two species were hybridized with probes derived from the known A, B, and C genomes; thus *S. pumila* and *S. parviflora* were designated as having genome D (Zhao et al. 2013). Similarly, the lack of hybridization signals with A, B, C, and D donor genomes led to the recognition of the E genome from *S. palmifolia* and the F genome from *S. arenaria* (Zhao et al. 2013). (Note that *S. arenaria* is a dubious name according to Morrone et al. (2014) and thus it is unclear what material was used by Zhao et al. (2013).) GISH also identified an apparent A genome autotetraploid, *S. apiculata* (=*queenslandica*). Analysis of *kn1* and *5S rDNA* sequences was consistent with the GISH results (Zhao et al. 2013).

The tetraploid *S. faberi* formed from an A genome (*S. viridis*) plus another genome from an unknown source closely related to *S. viridis* (Layton and Kellogg 2014). *S. faberi* is morphologically similar to *S. viridis*, but in the former species the upper glume is slightly shorter, so that the upper 1/4–1/3 of the upper lemma is visible (Layton and Kellogg 2014). In contrast, the upper glume of *S. viridis* completely covers the upper lemma and often slightly overlaps the lower lemma at the apex of the spikelet. *S. faberi* also has macrohairs on the adaxial surface of the leaves, whereas the leaves of *S. viridis* are glabrous. In addition, *S. faberi* is less tolerant of drought than *S. viridis* is and often grows in more mesic habitats such as the margins of cultivated fields, whereas *S. viridis* occurs more frequently in poor soil and cracks in pavement (Layton and Kellogg 2014).

Tetraploids classified as *S. verticillata* and *S. verticilliformis* each have one genome from the diploid *S. adhaerens* and one from *S. viridis* (Benabdelmouna et al. 2001b). Phylogenetic data using the low-copy nuclear marker *knotted1* on the same plant accessions confirmed the cytogenetic results, showing that *S. verticillata* and *S. verticilliformis* each have two loci, consistent with their ploidy level (Layton and Kellogg 2014). One *kn1* locus is related to that of the diploid *S. adhaerens* and the other to that of the diploid S. *viridis*.

The phylogenetic and cytogenetic data settle confusion over the taxonomy of *S. verticillata*, *S. adhaerens*, and *S. verticilliformis*. Some authors have considered the three species as synonymous (Rominger 1962; Doust et al. 2007; Kellogg et al. 2009), others have distinguished *S. verticillata* and *S. verticilliformis* but synonymized *S. adhaerens* with *S. verticillata* (Morrone et al. 2014), and still others maintain all three species as distinct (Rominger 2003). Both *S. verticillata* s.s. and *S. adhaerens* have retrorse prickles on the bristles, a character that is distinctive within

the genus and easily observed in the field because it makes the inflorescence adhere to clothing. Because this morphological character appears in the diploid and its derived polyploid, we can infer that it must be controlled by the B genome. *S. adhaerens* is distinctive in being a small plant with glabrous sheath margins and strigose hairs with papillose bases on the abaxial leaf surfaces. *S. verticillata*, in contrast, has sheath margins that are ciliate distally and abaxial leaf surfaces that are scabrous (Layton and Kellogg 2014). Some diploid chromosome numbers are reported for *S. verticillata* (Table 1.1) (Rominger 2003), but these are likely to be from misidentified specimens of *S. adhaerens* (Layton and Kellogg 2014).

*S. verticilliformis*, in contrast, has antrorse prickles but otherwise is morphologically indistinguishable from *S. verticillata*. Because of the genomic and morphological similarity of the two species, Layton and Kellogg (2014) recommend that the former species be placed in synonymy with the latter. They also recommend that *S. adhaerens* be recognized as a separate entity.

## 1.3 Summary

All organisms contain evidence of their evolutionary history and *Setaria* is no exception. Like all commelinid monocots, species of *Setaria* have extensive endosperm in their seeds, cell walls with glucuronoarabinoxylans, lignin with p-coumaric and ferulic acid, and leaves that accumulate silica. Like all Poales, the leaf silica accumulates in the epidermis, and like members of the graminid clade, the pollen is monoporate with a raised ring around the pore. *Setaria* has dumbbell shaped stomatal guard cells, and alternating long and short cells in the epidermis.

Like all other grasses, *Setaria* has a highly differentiated embryo, including a large haustorial derivative of the cotyledon, the scutellum. The fruit is a caryopsis in which the seed coat is fused to the inner epidermis of the pericarp. The antipodals proliferate in the female gametophyte. Short cells of the leaf epidermis produce bicellular microhairs. The genome has undergone a whole genome duplication dating to the origin of the family in the late Cretaceous. Like most but not quite all grasses, *Setaria* has flowers in spikelets with more than one flower per spikelet and two style branches and stigmas.

Like most other panicoid grasses (i.e., those in tribes Paniceae, Paspaleae, and Andropogoneae), *Setaria* has spikelets with exactly two flowers, the upper one bisexual and the lower one staminate or sterile. Spikelets are dorsiventrally compressed and develop basipetally. Leaf silica bodies are bilobed in surface view and symmetrical in cross section. Along with other members of tribe Paniceae, *Setaria* has a chromosome base number of $x=9$. Members of subtribe Cenchrinae share $C_4$ photosynthesis of the NADP-ME subtype and a single sheath surrounding the vascular bundle. The inflorescence bears sterile branches or bristles.

The genus *Setaria* itself has no diagnostic characters and is probably para- or polyphyletic as currently defined. Phylogenetic data are insufficient and not conclusive. Major clades correspond to geography and do not correlate well with morphol-

ogy. The Asian clade includes the model species *S. viridis*, along with its domesticated derivative *S. italica*, and its polyploid descendants *S. faberi*, *S. verticillata*, and *S. verticilliformis*.

## References


Aliscioni SS, Giussani LM, Zuloaga FO, Kellogg EA. A molecular phylogeny of *Panicum* (Poaceae: Paniceae): tests of monophyly and phylogenetic placement within the Paniceae. Am J Bot. 2003;90:796–821.

Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc. 2009;161:105–21.

Anton AM, Cocucci AE. The grass megagametophyte and its possible phylogenetic implications. Plant Syst Evol. 1984;146:117–21.

Baker G, Jones LHP, Wardrop ID. Cause of wear in sheep's teeth. Nature. 1959;184:1583–4.

Bala S, Sachdeva SK. Biochemical variability at cytotype level in *Setaria verticillata* (L.) Beauv. complex. Aspects Pl Sci. 1989;11:483–93.

Bala S, Sachdeva SK. Cytological and biochemical study of co-occurring diploid, tetraploid and hexaploid individuals of *Setaria verticillata* (L.) Beauv. (Poaceae). Phytologia. 1990;68:276–92.

Baltisberger M. Chromosomenzahlen einiger Pflanzen aus Albanien. II. Ber Geobot Inst Eidg Tech Hochsch Stift Rübel Zür. 1988;54:42–50.

Belaeva VA, Siplivinsky VN. In: Löve A, editor. Chromosome number reports LXXIII. Taxon, vol. 30; 1981. pp. 857–860

Benabdelmouna A, Abirached-Darmency M, Darmency H. Phylogenetic and genomic relationships in *Setaria italica* and its close relatives based on the molecular diversity 5S and 18S-5.8S-25S rDNA genes. Theor Appl Genet. 2001a;103:668–77.

Benabdelmouna A, Shi Y, Abirached-Darmency M, Darmency H. Genomic in situ hybridization (GISH) discriminates between the A and B genomes in diploid and tetraploid *Setaria* species. Genome. 2001b;44:685–90.

Bess EC, Doust AN, Kellogg EA. A naked grass in the "bristle clade": a phylogenetic and developmental study of *Panicum* section Bulbosa (Paniceae: Poaceae). Int J Plant Sci. 2005;166:371–81.

Bess EC, Doust AN, Davidse G, Kellogg EA. *Zuloagaea*, a new genus of tropical grass within the "bristle clade" (Poaceae: Paniceae). Syst Bot. 2006;31:656–70.

Bir SS, Chauhan HS. SOCGI plant chromosome number reports. IX. J Cytol Genet. 1990;25:147–8.

Bir SS, Sahni M. SOCGI plant chromosome number reports. I. J Cytol Genet. 1983;18:58–9.

Bir SS, Sahni M. SOCGI plant chromosome number reports. IV. J Cytol Genet. 1986;21:152–4.

Bir SS, Sahni M. Chromosomal and morphological variations in grasses of Punjab. J Cytol Genet. 1987;22:12–22.

Bowden WM, Seen HA. Chromosome numbers in 28 grass genera from South America. Can J Bot. 1962;40:1115–24.

Brown R. Prodromus florae Novae Hollandiae. London: J. Johnson; 1810.

Brown R. A voyage to Terra Australis. London: G. & W. Nicol; 1814.

Brown WV. A cytological study in the Gramineae. Am J Bot. 1948;35:382–95.

Campbell CS, Kellogg EA. Sister group relationships of the Poaceae. In: Soderstrom TR, Hilu KW, Campbell CS, Barkworth ME, editors. Grass systematics and evolution. Washington, DC: Smithsonian Institution; 1987.

Caponio I, Pensiero JF. Comportamiento citológico y reproductivo de *Setaria pflanzii* (Poaceae). Darwiniana. 2002;40:17–23.

Carpita NC. Structure and biogenesis of the cell walls of grasses. Annu Rev Plant Physiol Plant Mol Biol. 1996;47:445–76.

Chemisquy MA, Giussani LM, Scataglini M, Kellogg EA, Morrone O. Phylogenetic studies favor the unification of *Pennisetum*, *Cenchrus* and *Odontelytrum*: a nuclear, chloroplast and morphological combined analysis. Ann Bot. 2010;106:107–30.

Chikara J, Gupta PK. Karyological studies in the genus *Setaria.* I. Variability within *Setaria italica* (L.) Beauv. J Cytol Genet. 1979;14:75–9.
Chopanov P, Yurtsev BN. Chromosome numbers of some grasses of Turkmenia. II. Bot Zhurn SSSR. 1976;61:1240–4.
Christopher J, Abraham A. Studies on the cytology and phylogeny of South Indian grasses. III. Subfamily VI: Panicoideae, tribe Paniceae. Cytologia. 1976;41:621–37.
Clayton WD, Harman KT, Williamson H. GrassBase—the online world grass flora [Online]. 2006 onwards. http://www.kew.org/data/grasses-db.html (2010).
Darmency H, Pernes J. An inheritance study of domestication in foxtail millet using an interspecific cross. Plant Breed. 1987;99:30–3.
Darmency H, Zangre GR, Pernes J. The wild-weed-crop complex in *Setaria*: a hybridization study. Genetica. 1987;75:103–7.
Devesa JA, Ruiz T, Viera MC, Tormo R, Vázquez F, Carrasco JP, Ortega A, Pastor J. Contribución al conocimiento cariológico de las Poaceae en Extremadura (España) III. Bol Soc Broteriana, Ser 2. 1991;64:35–74.
deWet JMJ. Chromosome numbers of a few South African grasses. Cytologia. 1954;19:97–103.
deWet JMJ. Additional chromosome numbers in Transvaal grasses. Cytologia. 1958;23:113–8.
deWet JMJ, Anderson LJ. Chromosome numbers in Transvaal grasses. Cytologia. 1956;21:1–10.
Donadio S, Giussani LM, Kellogg EA, Zuloaga FO, Morrone O. A molecular phylogeny of *Pennisetum* and *Cenchrus* (Poaceae-Paniceae) based on the *trnL-F*, *rpl16* chloroplast markers. Taxon. 2009;58:392–404.
Doust AN, Kellogg EA. Inflorescence diversification in the panicoid "bristle grass" clade (Paniceae, Poaceae): evidence from molecular phylogenies and developmental morphology. Am J Bot. 2002;89:1203–22.
Doust AN, Penly AM, Jacobs SWL, Kellogg EA. Congruence, conflict and polyploidization shown by nuclear and chloroplast markers in the monophyletic "bristle clade" (Paniceae, Panicoideae, Poaceae). Syst Bot. 2007;32:531–44.
Dujardin M. Chromosome numbers of some tropical African grasses from western Zaire. Can J Bot. 1978;56:2138–52.
Emery WHP. A cyto-taxonomic study of *Setaria macrostachya* (Gramineae) and its relatives in the southwestern United States and Mexico. Bull Torrey Bot Club. 1957a;84:94–105.
Emery WHP. A study of reproduction in *Setaria macrostachya* and its relatives in the southwestern United States and northern Mexico. Bull Torrey Bot Club. 1957b;84:106–21.
Evans MMS, Grossniklaus U. The maize megagametophyte. In: Bennetzen JL, Hake SC, editors. Handbook of maize: its biology. Berlin: Springer; 2009.
Fernández A, Queiróz M. Contribution à la connaissance cytotaxonomique des spermatophita du Portugal. Bol Soc Broteriana, Ser 2. 1969;18:3–140.
Franks PJ, Farquhar GD. The mechanical diversity of stomata and its significance in gas-exchange control. Plant Physiol. 2007;143:78–87.
Frey L, Mizianty M, Mirek Z. Chromosome numbers of Polish vascular plants. Fragm Florist Geobot. 1981;27:581–90.
Gadella TWJ. Chromosome number reports LVI. Taxon. 1977;26:257–74.
Garbuzov M, Reidinger S, Hartley SE. Interactive effects of plant-available soil silicon and herbivory on competition between two grass species. Ann Bot. 2011;108:1355–63.
Giussani LM, Cota-Sánchez JH, Zuloaga FO, Kellogg EA. A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of $C_4$ photosynthesis. Am J Bot. 2001;88:1993–2012.
Givnish TJ, Ames M, Mcneal JR, Mc Kain MR, Steele PR, de Pamphilis CW, Graham SW, Pires JC, Stevenson DW, Zomlefer WB, Briggs BG, Duvall MR, Moore MJ, Heaney JM, Soltis DE, Soltis PS, Thiele K, Leebens-Mack JH. Assembling the tree of monocotyledons: plastome sequence phylogeny and evolution of Poales. Ann Mo Bot Gard. 2010;97:584–616.
Goff SA, Ricke D, Lan T-H, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y,

Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun W-L, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science. 2002;296:92–100.

Gómez-Martinez R, Culham A. Phylogeny of the subfamiily Panicoideae with emphasis on the tribe Paniceae: evidence from the *trnL-F* cp DNA region. In: Jacobs SWL, Everett J, editors. Grasses: systematics and evolution. Collingwood: Commonwealth Scientific and Industrial Research Organization (CSIRO); 2000.

Gould FW, Soderstrom TR. Chromosome numbers of tropical American grasses. Am J Bot. 1967;54:676–83.

Grass Phylogeny Working Group (GPWG). Phylogeny and subfamilial classification of the Poaceae. Ann Mo Bot Gard. 2001;88:373–457.

Grass Phylogeny Working Group II (GPWG II). New grass phylogeny resolves deep evolutionary relationships and discovers $C_4$ origins. New Phytol. 2012;193:304–12.

Gupta PK, Singh RV. Variations in chromosomes and flavonoids in *Setaria* Beauv. Nucleus. 1977;20:167–71.

Guzik MB. Khromosomnye chisla nekotorykh dikorastushchkh zlakov Predural'ja. Ekologija Opylenija Rastenij. 1984;8:82–6.

Harris PJ, Hartley RD. Phenolic constituents of the cell walls of monocotyledons. Biochem Syst Ecol. 1980;8:153–60.

Harris PJ, Tretheway JAK. The distribution of ester-linked ferulic acid in the cell walls of angiosperms. Phytochem Rev. 2009;9:19–33.

Hattersley PW. Variations in photosynthetic pathway. In: Soderstrom TR, Hilu KW, Campbell CS, Barkworth ME, editors. Grass systematics and evolution. Washington, DC: Smithsonian Institution Press; 1987. p. 49–64.

Hattersley PW, Watson L. Diversification of photosynthesis. In: Chapman GP, editor. Grass evolution and domestication. Cambridge: Cambridge University Press; 1992.

Haworth M, Elliott-Kingston C, Mc Elwain JC. Stomatal control as a driver of plant evolution. J Exp Bot. 2011;62:2419–23.

Hirano R, Naito K, Fukunaga K, Watanabe KN, Ohsawa R, Kawase M, Belzile F. Genetic structure of landraces in foxtail millet (*Setaria italic*a (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet. Genome. 2011;54:498–506.

Huang P, Feldman M, Schroder S, Bahri BA, Diao X, Zhi H, Estep M, Baxter I, Devos KM, Kellogg EA. Population genetics of *Setaria viridis*, a new model system. Mol Ecol. 2014;23:4192–295.

Hunt HV, Van Der Linden M, Liu X, Motuzaite-Matuzeviciute G, Colledge S, Jones MK. Millets across Eurasia: chronology and context of early records of the genera *Panicum* and *Setaria* from archaeological sites in the Old World. Veg Hist Archaeobot. 2008;17(suppl):S5–18.

Isa M, Bai S, Yokoyama T, Ma JF, Ishibashi Y, Yuasa T, Iwaya-Inoue M. Silicon enhances growth independent of silica deposition in a low-silica rice mutant, *lsi1*. Plant Soil. 2010;331:361–75.

Johnston CR, Watson L. Microhairs: a universal characteristic of non-festucoid grass genera? Phytomorphology. 1976;26:297–301.

Kellogg EA. The grasses: a case study in macroevolution. Ann Rev Ecol Syst. 2000;31:217–38.

Kellogg EA. Poaceae. In: Kubitzki K, editor. The families and genera of vascular plants. Berlin: Springer; 2015. pp. 1–416.

Kellogg EA, Campbell CS. Phylogenetic analyses of the Gramineae. In: Soderstrom TR, Hilu KW, Campbell CS, Barkworth ME, editors. Grass systematics and evolution. Washington, DC: Smithsonian Institution Press; 1987. p. 310–22.

Kellogg EA, Linder HP. Phylogeny of Poales. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ, editors. Monocotyledons: systematics and evolution. Kew: Royal Botanic Gardens; 1995.

Kellogg EA, Hiser KM, Doust AN. Taxonomy, phylogeny, and inflorescence development of the genus *Ixophorus* (Panicoideae: Poaceae). Int J Plant Sci. 2004;165:1089–105.
Kellogg EA, Aliscioni SS, Morrone O, Pensiero J, Zuloaga FO. A phylogeny of *Setaria* (Poaceae, Panicoideae, Paniceae) and related genera, based on the chloroplast gene *ndhF*. Int J Plant Sci. 2009;170:117–31.
Kellogg EA, Camara PEAS, Rudall PJ, Ladd P, Malcomber ST, Whipple CJ, Doust AN. Early inflorescence development in the grasses (Poaceae). Front Plant Sci. 2013;4:250. doi:10.3389/fpls.2013.00250.
Khosla PK, Sharma ML. Cytological observations on some species of *Setaria*. Nucleus. 1973;26:38–41.
Kliphuis E, Wieffering JH. IOPB chromosome number reports LXIV. Taxon. 1979;28:398–400.
Koul KK, Gohil RN. SOCGI plant chromosome number reports. VI. J Cytol Genet. 1988;23:38–52.
Koul KK, Gohil RN. Cytogenetic studies on some Kashmir grasses. VIII. Tribe Agrostideae, Festuceae and Paniceae. Cytologia. 1991;56:437–52.
Kozuharov SI, Petrova AV. Chromosome numbers of Bulgarian angiosperms. Fitologija. 1991;39:72–7.
Lakshmi N, Yacob Z. A new chromosome number report in *Setaria kagerensis* Mez. Curr Sci. 1978;47:813.
Layton DJ, Kellogg EA. Morphological, phylogenetic, and ecological diversity of the new model species *Setaria viridis* (Poaceae: Paniceae) and its close relatives. Am J Bot. 2014;101:539–57.
Le Thierry d'Ennequin M, Panaud O, Brown S, Siljak-Yakoviev S, Sarr A. First evaluation of nuclear DNA content in *Setaria* genus by flow cytometry. J Hered. 1998;89:556–9.
Le Thierry d'Ennequin M, Panaud O, Toupance B, Sarr A. Assessment of genetic relationships between *Setaria italica* and its wild relative *S. viridis* using AFLP markers. Theor Appl Genet. 2000;100:1061–6.
Li X, Chen R. The karyotype analysis of *Setaria italica* and *S. viridis*. J Wuhan Bot Res. 1985;3:409–12.
Li HW, Li CH, Pao WK. Cytological and genetical studies of the interspecific cross of the cultivated foxtail millet, *Setaria italica* (L.) Beauv., and the green foxtail millet, *S. viridis* L. J Am Soc Agron. 1945;37:32–54.
Li M, Song YC, Liu H. Studies on G-banding karyotypes in millet. J Henan Vocat Teach Coll. 1996;24:1–7.
Linder HP, Rudall PJ. Evolutionary history of Poales. Ann Rev Ecol Evol Syst. 2005;36:107–24.
Löve A, Löve D. Chromosome number reports LXXI. Taxon. 1981;30:509–11.
Ma JF, Yamaji N. Silicon uptake and accumulation in higher plants. Trends Pl Sci. 2006;11:392–7.
Magulaev AJ. The chromosome numbers of flowering plants of the northern Caucasus. Pt. II. In: Galushko AI, editor. Flora of the northern Caucasus, vol. 2. Rostov: Rostovskiy Universitet; 1976.
Malcomber ST, Kellogg EA. Heterogeneous expression patterns and separate roles of the SEPALLATA gene *Leafy Hull Sterile* (*LHS1*) in grasses. Plant Cell. 2004;16:1692–706.
Manero de Zamelzú D, Ochoa de Suárez B. Número cromosómico de *Setaria leiantha* (Poaceae). Agriscientia. 1991;8:79.
Massey FP, Hartley SE. Physical defences wear you down: progressive and irreversible impacts of silica on insect herbivores. J Anim Ecol. 2009;78:281–91.
McKain MR, Tang H, McNeal JR, Ayyampalayam S, Davis JI, dePamphilis CW, Givnish TJ, Pires JC, Stevenson DW, Leebens-Mack JH. A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. Genome Biol Evol. 2016;8:1150–64.
Mehra PN. Cytology of East Indian grasses. India: Chandigarh; 1982.
Mehra PN, Sharma ML. Cytological studies in some central and eastern Himalayan grasses. II. The Paniceae. Cytologia. 1975;40:75–89.
Moffett AA, Hurcombe R. Chromosome numbers of South African grasses. Heredity. 1949; 3:369–73.

Molinari HB, Pellny TK, Freeman J, Shewry PR, Mitchell RA. Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in *Brachypodium distachyon*. Front Plant Sci. 2013;4:50.

Morrone O, Aagesen L, Scataglini MA, Salariato DL, Denham SS, Chemisquy MA, Sede SM, Giussani LM, Kellogg EA, Zuloaga FO. Phylogeny of the Paniceae (Poaceae: Panicoideae): integrating plastid DNA sequences and morphology into a new classification. Cladistics. 2012;28:333–56.

Morrone O, Aliscioni SA, Veldkamp JF, Pensiero J, Zuloaga FO, Kellogg EA. A revision of the Old World species of *Setaria*. Syst Bot Monogr. 2014;96:1–161.

Mulligan GA. Chromosome numbers of some plants native and naturalized in Canada. Nat Can. 1984;111:447–9.

Nadeem Ahsan SM, Vahidy AA, Ali SI. Chromosome numbers and incidence of polyploidy in Panicoideae (Poaceae) from Pakistan. Ann Mo Bot Gard. 1994;81:775–83.

Nicola MV, Lizarazu MA, Scataglini MA. Phylogenetic analysis and taxonomic position of section *Verrucosa* of *Panicum* and its relationship with taxa of the *Sacciolepis-Trichanthecium* clade (Poaceae: Panicoideae: Paniceae). Plant Syst Evol. 2015;301:2247.

Norrmann GA, Quarín CL, Killeen TJ. Chromosome numbers in Bolivian grasses (Gramineae). Ann Mo Bot Gard. 1994;81:768–74.

Oliveira Freitas-Sacchet AMD. Chromosome number reports LXIX. Taxon. 1980;29:703–4.

Oliveira Freitas-Sacchet AMD, Boldrini II, Born GG. Cytogenetics and evolution of the native grasses of Rio Grande do Sul, Brazil, *Setaria* Beauv. (Gramineae). Rev Bras Genet. 1984;7:535–48.

Olorode O. Additional chromosome counts in Nigerian grasses. Brittonia. 1975;27:63–8.

Paterson AH, Bowers JE, Chapman BA. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A. 2004;101:9903–8.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob UR, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS. The *Sorghum bicolor* genome and the diversification of grasses. Nature. 2009;457:551–6.

Pensiero JF. Las especies sudamericanas del género *Setaria* (Poaceae, Paniceae). Darwiniana. 1999;37:37–151.

Petrik DL, Karlen SD, Cass CL, Padmakshan D, Lu F, Liu S, Le Bris P, Antelme S, Santoro N, Wilkerson CG, Sibout R, Lapierre C, Ralph J, Sedbrook JC. p-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in *Brachypodium distachyon*. Plant J. 2014;77:713–26.

Piperno DR. Phytoliths: a comprehensive guide for archaeologists and paleoecologists. New York: AltaMira; 2006.

Piperno DR, Pearsall DM. The silica bodies of tropical American grasses: morphology, taxonomy, and implications for grass systematics and fossil phytolith identification. Smithson Contrib Bot. 1998;85:1–40.

Pohl RW, Davidse G. Chromosome numbers of Costa Rican grasses. Brittonia. 1971;23:293–324.

Prendergast HDV, Hattersley PW. Australian $C_4$ grasses (Poaceae): leaf blade anatomical features in relation to $C_4$ acid decarboxylation types. Aust J Bot. 1987;35:355–82.

Prendergast HDV, Hattersley PW, Stone NE. New structural/biochemical associations in leaf blades of $C_4$ grasses (Poaceae). Aust J Plant Physiol. 1987;14:403–20.

Preston JC, Christensen A, Malcomber ST, Kellogg EA. MADS-box gene expression and implications for developmental origins of the grass spikelet. Am J Bot. 2009;96:1419–29.

Probatova NS, Sokolovskaya AP. Chromosome numbers in Adoxaceae, Chloranthaceae, Cupressaceae, Juncaceae, Poaceae. Bot Zhurn SSSR. 1983;68:1683.

Quarín CL. Recuentos cromosómicos en gramíneas de Argentina subtropical. Hickenia. 1977;1:73–8.
Raman VS, Chandrasekharan P, Krishnaswami D. Chromosome numbers in Gramineae. Curr Sci. 1959;28:453–4.
Rao PN, Mwasumbi LB. Chromosome number reports LXX. Taxon. 1981;30:79–80.
Reeder JR. Notes on Mexican grasses. IX. Miscellaneous chromosome numbers 3. Brittonia. 1971;23:105–17.
Rominger JMD. Taxonomy of *Setaria* (Gramineae) in North America. Illinois Biol Monogr. 1962;29:1–132.
Rominger JMD. *Setaria* P. Beauv. In: Barkworth ME, Capels KM, Long S, Piep MB, editors. Flora of North America North of Mexico, Magnoliophyta: Commelinidae (in part): Poaceae, part 1, vol. 25. New York: Oxford University Press; 2003.
Rudall PJ, Stuppy W, Cunniff J, Kellogg EA, Briggs BG. Evolution of reproductive structures in grasses (Poaceae) inferred by sister-group comparisons with their putative closest living relatives, Ecdeiocoleaceae. Am J Bot. 2005;92:1432–43.
Sahni M. Cytological variability in genus *Setaria* P. Beauv. from Punjab plains. Aspects Pl Sci. 1989;11:467–73.
Sahni M, Bir SS. SOCGI plants chromosome number reports—III. J Cytol Genet. 1985;20:205–6.
Sajo MG, Longhi-Wagner H, Rudall PJ. Floral development and embryology in the early-divergent grass *Pharus*. Int J Plant Sci. 2007;168:181–91.
Sajo MG, Longhi-Wagner HM, Rudall PJ. Reproductive morphology of the early-divergent grass *Streptochaeta* and its bearing on the homologies of the grass spikelet. Plant Syst Evol. 2008;275:245–55.
Sánchez-Ken JG, Clark LG. Phylogenetic relationships within the Centothecoideae + Panicoideae clade (Poaceae) based on *ndhF* and *rpl16* intron sequences and structural data. Aliso. 2007; 23:487–502.
Sánchez-Ken JG, Clark LG. Phylogeny and a new tribal classification of the Panicoideae s.l. (Poaceae) based on plastid and nuclear sequence data and structural data. Am J Bot. 2010;97:1732–48.
Sanson GD, Kerr SA, Gross KA. Do silica phytoliths really wear mammalian teeth? J Archaeol Sci. 2007;34:526–31.
Sarkar AK, Datta N, Mallick R, Chatterjee U. In IOPB chromosome number reports LIV. Taxon. 1976;25:648–9.
Saxena BK, Gupta BK. Chromosome numbers of some grasses of Dehra Dun. Bull Bot Surv India. 1969;11:443–4.
Sede S, Escobar A, Morrone O, Zuloaga FO. Chromosome studies in American Paniceae (Poaceae, Panicoideae). Ann Mo Bot Gard. 2010;97:128–38.
Shadowsky AE. Der antipodale Apparat bei Gramineen. Flora. 1926;120:344–70.
Sharma ML, Sharma K. Cytological studies in the North Indian grasses. Cytologia. 1979;44:861–72.
Shi Y, Wang T, Li Y, Darmency H. Impact of transgene inheritance on the mitigation of gene flow between crops and their wild relatives: the example of foxtail millet. Genetics. 2008;180:969–75.
Simpson GG. Horses: the story of the horse family in the modern world and through sixty million years of history. Oxford: Oxford University Press; 1951.
Singh DN, Godward MBE. Cytological studies in the Gramineae. Heredity. 1960;15:193–9.
Singh RV, Gupta PK. Cytological studies in the genus *Setaria* (Gramineae). Cytologia. 1977;42:483–93.
Sinha RRP, Bhardwaj AK, Singh RK. SOCGI plant chromosome number reports. IX. J Cytol Genet. 1990;25:140–3.
Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, Filgueiras TS, Davis JI, Morrone O. A worldwide phylogenetic classification of the Poaceae (Gramineae). J Syst Evol. 2015;53:117–37.
Spies JJ, Du Plessis H. Chromosome studies on African plants. 1. Bothalia. 1986;16:87–8.

Stevens PF. Angiosperm phylogeny website. Version 12, July 2012 [and more or less continuously updated since]. 2001 onwards. [Online]. 2012. http://www.mobot.org/MOBOT/research/APweb/

Strömberg CAE. Evolution of hypsodonty in equids: testing a hypothesis of adaptation. Paleobiology. 2006;32:236–58.

Sylvester AW, Parker-Clark V, Murray GA. Leaf shape and anatomy as indicators of phase change in the grasses: comparison of maize, rice and bluegrass. Am J Bot. 2001;88:2157–67.

Takacs EM, Li J, Du C, Ponnala L, Janick-Buckner D, Yu J, Muehlbauer GJ, Schnable PS, Timmermans MCP, Sun Q, Nettleton D, Scanlon MJ. Ontogeny of the maize shoot apical meristem. Plant Cell. 2012;24:3219–34.

Tiranti IN, Genghini RN. Caracteres cuantitativos de la microsporogénesis de cuatro especies de *Setaria* (Poaceae). Bol Soc Argent Bot. 2000;35:331–3.

Váchová M. Index to chromosome numbers of Slovakian flora. Pt. 6. Acta Fac Rerum Nat Univ Comen Bot. 1978;26:1–42.

Váchová M, Feráková V. Chromosome number reports LXIX. Taxon. 1980;29:722–3.

Veldkamp JF. Miscellaneous notes on southeast Asian Gramineae. IX. *Setaria* and *Paspalidium*. Blumea. 1994;39:373–84.

Vicentini A, Barber JC, Giussani LM, Aliscioni SS, Kellogg EA. Multiple coincident origins of $C_4$ photosynthesis in the Mid- to Late Miocene. Glob Chang Biol. 2008;14:2963–77.

Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD. Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1998;96:31–6.

Wang Y, Zhi H, Li W, Li H, Wang Y, Huang Z, Diao X. A novel genome of C and the first autotetraploid species in the *Setaria* genus identified by genomic in situ hybridization (GISH). Genet Res Crop Evol. 2009;56:843–50.

Warwick SI, Thompson BK, Black LD. Life history and allozyme variation in populations of the weed species *Setaria faberi*. Can J Bot. 1987;65:1396–402.

Warwick SI, Crompton C, Black LD, Wotjas W. IOPB chromosome data 11. Newsl Int Organ Plant Biosyst. 1997;26/27:25–6.

Washburn JD, Schnable JC, Davidse G, Pires JC. Phylogeny and photosynthesis of the grass tribe Paniceae. Am J Bot. 2015;102:1493–505.

Webster RD. Nomenclature of *Setaria* (Poaceae: Paniceae). Sida. 1993;15:447–89.

Webster RD. Nomenclatural changes in *Setaria* and *Paspalidium* (Poaceae: Paniceae). Sida. 1995;16:439–46.

Withers S, Lu F, Kim H, Zhu Y, Ralph J, Wilkerson CG. Identification of grass-specific enzyme that acylates monolignols with p-coumarate. J Biol Chem. 2012;287:8347–55.

Wu QM, Bai JL. Cytogenetic and isoenzymic studies on *Setaria* millet and *S. verticillata* (2X) and *S. verticiformis* (4X). Acta Botan Boreali Occiden Sin. 2000;20:954–9.

Xu BS, Weng RF, Zhang MZ. Chromosome numbers of of Shanghai plants I. Investig Studium Nat. 1992;12:48–65.

Yu J, Hu S, Wang J, Wong GK-S, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science. 2002;296:79–92.

Zhao M, Zhi H, Doust AN, Li W, Wang Y, Li H, Jia G, Wang Y, Zhang N, Diao X. Novel genomes and genome constitutions identified by GISH and 5S rDNA and *knotted1* genomic sequences in the genus *Setaria*. BMC Genomics. 2013;14:244.

Zhou X, Sun PY, Chou YL, Hou BY, Sun T. A study on the basic karyotype of *Setaria italica*. Sci Agric Sin. 1989;22:30–4.

Zuloaga FO, Morrone O, Scataglini MA. Monograph of *Trichanthecium* (Poaceae, Paniceae). Syst Bot Monogr. 2011;94:1–101.

# Chapter 2
# Population Genetics and Genome-Wide Association Mapping of Chinese Populations of Foxtail Millet and Green Foxtail

**Guanqing Jia**

**Abstract** Green foxtail (*Setaria viridis*) and foxtail millet (*Setaria italica*) are new biological and genomic models for investigation of the biology of $C_4$ photosynthesis and grass evolution. Green foxtail is the ancestor of foxtail millet, an ancient cereal of great importance in arid and semi-arid regions of the world, especially in China and India. To date, China has been recognized as the center of origin and improvement of foxtail millet, and over 80% of the world's Setaria accessions are conserved in the National Gene Bank of China. Assessment of germplasm samples collected in China can help to reveal the domestication history and potential for improvement of cultivated foxtail millet. Recently, the molecular diversity, genetic structure, eco-geographical distribution and selection history of foxtail millet cultivars has been revealed through large scale germplasm characterization, genomic analysis, and genome-wide association mapping of QTLs controlling agronomic traits. These achievements have laid the foundation for further exploration of functional genes controlling vital characters in Setaria and will be powerful tools for improved marker-assisted breeding of foxtail millet cultivars. In this chapter, recent studies on the Chinese Setaria gene pool will be discussed, as well as their potential for benefiting future genetic investigations in Setaria.

**Keywords** Setaria • Germplasm • Diversity • Structure • Domestication • Breeding • Genome-wide association studies (GWAS) • Quantitative trait loci (QTL)

## 2.1 Introduction

According to recent archeological evidence, foxtail millet (*Setaria italica*) has been domesticated and grown for harvesting grains as food in North China for over 11,500 years (Lu et al. 2009; Yang et al. 2012). As one of the oldest cereals cultivated

G. Jia, Ph.D. (✉)
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences,
No. 12 Zhongguancun South St., Haidian District, Beijing 100081,
People's Republic of China
e-mail: jiaguanqing@caas.cn



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_2

widely in China, large numbers of foxtail millet variants have emerged since domestication. To date, over 27,500 accessions of foxtail millet collected from China have been conserved in the National Gene Bank of China, operated by the Chinese Academy of Agricultural Sciences (CAAS). China is also the center of foxtail millet crop production and contributes over 80 % of the global grain yield of this ancient cereal. For a long time, China has been the main region where most foxtail millet improvement has been conducted, despite the fact that basic research in Setaria has been carried out recently in a number of countries across the world (Doust et al. 2009; Li and Brutnell 2011; Diao et al. 2014).

Evaluation and identification of genetic diversity and population structure of germplasm accessions is pivotal for resource management and hybrid breeding of foxtail millet. In addition, association mapping of Quantitative-Trait-Loci (QTLs) controlling agronomically important traits through Genome-Wide-Association-Studies (GWAS) can provide more opportunities for breeders to characterize genetic backgrounds and to pyramid favorable traits by molecular marker approaches.

This chapter describes the diversity and structure of Chinese Setaria accessions, including landraces, improved cultivars, and wild relatives, together with an overview of molecular and genetic analyses to identify genomic regions contributing to important morphological traits in foxtail millet. Finally, conclusions and research perspectives will be presented, especially in the context of the rapid development of a new generation of high-throughput genome re-sequencing technologies.

## 2.2 Diversity and Structure of Chinese Setaria Accessions

The first-generation haplotype map of genomic variants of foxtail millet was constructed through genome re-sequencing of 916 accessions of foxtail millet lines (Jia et al. 2013a). Sequence diversity ($\pi$) of cultivated Setaria accessions was estimated to be ~0.0010, which is comparable to the diversity of *indica* (~0.0016) and *japonica* (~0.0006) subspecies of rice (*Oryza sativa*). The population structure of Chinese Setaria accessions (including both landraces and improved elite cultivars) is highly consistent with the geographic distribution of eco-regions in China. Two divergent subgroups, Part 1 and Part 2, which include accessions sourced from north China and south China respectively, are inferred by phylogenic analysis (Jia et al. 2013a). Linkage disequilibrium (LD) decay rate in foxtail millet is ~100 Kb on average, similar to cultivated rice (Huang et al. 2010) and sorghum (*Sorghum bicolor*) (Morris et al. 2013), which are also self-pollinated organisms.

### *2.2.1 Landraces*

A landrace is a kind of locally adapted variety domesticated from wild species of animals or plants, by long-term human or environmental selection and population isolation. Landraces usually display a high level of diversity in morphological

phenotypes, physiological characters, agronomic traits, and even genomic sequences. In China, prehistoric remains chart the long history of foxtail millet domestication. In northern China, the domestication of foxtail millet was initiated 11,500 years ago and was the predominant food crop until 4,000 years ago (Lu et al. 2005). The area under cultivation continued to increase up until the 1950s. There exist ancient agricultural documents that list foxtail millet as an important cereal food crop, and several officials in the Qing and Ming dynasties in China were designated as "officers of foxtail millet management" (Cao 1986).

The diversity of foxtail millet landraces is the result of conscious selection by farmers for specific phenotypes for harvesting and sowing, including (1) Higher yield potential; (2) Better biotic and abiotic stress tolerance, good adaptability, and lodging resistance; (3) Suitability for local farming environmental conditions and for intercropping; and (4) Good quality of kernel/grain ratio (ratio of milled grain to harvested grain), and better nutrition and flavor (Li and Wu 1996).

According to a morphological analysis of accessions conserved in the Chinese national gene bank reported by Li et al. (1995), foxtail millet landraces sourced from different eco-regions of China are very diverse. Landraces differed in all measured morphological traits (Table 2.1). Overall, grain dry weight and number of tillers were more variable (exhibited a higher Coefficient of Variation, CV) than other traits. For landraces collected from other countries, grain dry weight and number of tillers also exhibited a higher level of phenotypic variation. In terms of yield-related characters, Chinese landraces produced a higher level of grain yield in field production, but crude protein contents of Chinese landraces were lower than traditional cultivars sampled from other countries. Reddy et al. (2006) reported on the morphological diversity of foxtail millet germplasm collections conserved at the International Crops Research Institute for Semi-Arid Tropics (ICRISAT) in India and showed that there was more variation among landraces sampled from China than among landraces sampled from any other country. Overall, however, variation among landraces outside China was greater than that within China (Table 2.1), which may be due to fewer number of samples collected and conserved in other countries (Li et al. 1995).

The genetic diversity of foxtail millet landraces was analyzed by Jusuf and Pernes (1985) using ten isozymic polymorphic loci. Their results suggested that China has the highest level of molecular diversity for the species. RAPD markers (Li et al. 1998; Schontz and Rether 1999), AFLP markers (Le Thierry d'Ennequin et al. 2000), and ISSR markers (Li et al. 2012) have also been utilized for foxtail millet landrace diversity analysis. Jia et al. (2009) characterized genetic diversity of 40 foxtail millet landraces sourced from China using SSR markers. Van et al. (2008) detected the LD extent in Waxy gene sequences from a set of a worldwide collections of foxtail millet, and Wang et al. (2010) analyzed extent of LD in nine selected genomic fragments from 50 landrace accessions, with both studies revealing low levels of LD in foxtail millet. A report of population structure analysis by transposon display (TD) classified foxtail millet landraces into eight clusters that are linked closely to geography and suggest a monophyletic origin of foxtail millet domestication (Hirano et al. 2011). In recent studies reported by Wang et al. (2012), a total of 250 representative foxtail millet landraces collected from China were sampled for population genetic analysis using microsatellite markers. For each SSR locus, an average of 20.9 alleles

**Table 2.1** Mean values of quantitative traits of foxtail millet landraces sourced from China and other countries

| Province/ country | Till | Hgt (cm) | CulD (cm) | Node# | PanL (cm) | DaysM (day) | DWP (g) | DW1000 (g) | Prot (%) |
|---|---|---|---|---|---|---|---|---|---|
| Anhui | 1.5 | 154.1 | 0.50 | 16.0 | 20.0 | 107.0 | 11.5 | 2.68 | 13.00 |
| Beijing | 1.4 | 139.0 | 0.67 | 15.0 | 24.0 | 105.3 | 10.8 | 2.68 | 14.32 |
| Fujian | 1.0 | 170.9 | 0.60 | 18.2 | 27.9 | 124.0 | 14.5 | 1.10 | 11.45 |
| Guangdong | 2.4 | 123.7 | 0.66 | 16.7 | 24.8 | 129.0 | 5.2 | 1.57 | 16.04 |
| Guangxi | 1.1 | 140.3 | 0.70 | 23.4 | 21.7 | 112.8 | 3.8 | 2.14 | 13.20 |
| Guizhou | 1.8 | 147.3 | 0.67 | 16.0 | 23.3 | 119.7 | 7.9 | 2.25 | 16.22 |
| Hebei | 2.5 | 131.8 | 0.70 | 17.4 | 21.0 | 111.8 | 11.9 | 2.63 | 14.00 |
| Henan | 2.2 | 131.4 | 0.66 | 13.8 | 20.9 | 116.4 | 11.1 | 2.52 | 15.12 |
| Hunan | 2.2 | 129.4 | 0.70 | 16.9 | 23.5 | 127.0 | 10.4 | 1.63 | 15.10 |
| Jiangsu | 1.7 | 113.4 | 0.70 | 14.4 | 18.4 | 110.0 | 10.1 | 2.60 | 14.30 |
| Shaanxi | 1.3 | 138.8 | 0.75 | 14.4 | 24.9 | 123.3 | 9.1 | 2.71 | 15.42 |
| Shandong | 2.0 | 130.4 | 0.67 | 14.7 | 21.5 | 113.8 | 12.0 | 2.70 | 13.94 |
| Shanxi | 1.2 | 132.4 | 0.73 | 13.9 | 25.0 | 117.1 | 10.1 | 2.97 | 15.19 |
| Xinjiang | 2.4 | 113.8 | 0.47 | 12.3 | 27.0 | 95.4 | 5.3 | 2.86 | 17.36 |
| Yunnan | 1.4 | 122.8 | 0.42 | 12.8 | 18.7 | 98.7 | 5.3 | 1.57 | 15.04 |
| Zhejiang | 1.4 | 121.1 | 0.70 | 12.3 | 21.2 | 94.5 | 9.6 | 2.34 | 14.96 |
| Average | 1.7 | 133.8 | 0.64 | 15.5 | 22.7 | 112.9 | 9.3 | 2.31 | 14.67 |
| CV% | 29.22 | 11.10 | 15.00 | 17.75 | 12.19 | 9.47 | 32.30 | 24.00 | 9.56 |
| | | | | | | | | | |
| Afghanistan | 6.4 | 74.4 | 0.27 | 9.1 | 10.1 | 90.7 | 2.6 | 2.43 | 15.33 |
| Albania | 1.1 | 132.9 | 0.50 | 11.5 | 26.1 | 108.0 | 5.0 | 2.50 | 15.06 |
| Australia | 2.9 | 96.1 | 0.37 | 8.3 | 14.9 | 97.7 | 5.9 | 2.60 | 16.48 |
| Bulgaria | 2.6 | 121.3 | 0.40 | 11.1 | 16.6 | 105.0 | 5.0 | 2.02 | 19.58 |
| Denmark | 1.6 | 102.4 | 0.30 | 8.8 | 16.1 | 85.0 | 2.7 | 2.40 | 14.22 |
| Finland | 1.1 | 70.0 | 0.30 | 8.4 | 17.5 | 90.0 | 1.3 | 2.04 | 17.37 |
| France | 1.0 | 121.0 | 0.42 | 11.9 | 26.0 | 96.2 | 3.8 | 2.45 | 15.94 |
| Hungary | 2.1 | 110.1 | 0.48 | 10.1 | 13.6 | 94.3 | 3.6 | 2.55 | 17.79 |
| India | 4.4 | 137.0 | 0.53 | 16.7 | 17.1 | 120.0 | 6.5 | 2.29 | 16.54 |
| Iran | 4.3 | 83.6 | 0.30 | 10.8 | 12.5 | 90.0 | 2.4 | 2.40 | 15.78 |
| Italy | 1.0 | 132.0 | 0.50 | 12.5 | 24.5 | 111.0 | 8.0 | 2.88 | 17.27 |
| Japan | 1.4 | 136.3 | 0.68 | 15.3 | 23.6 | 114.4 | 11.3 | 2.23 | 14.66 |
| Kenya | 4.2 | 146.7 | 0.54 | 16.5 | 17.9 | 115.6 | 8.1 | 2.23 | 16.44 |
| Korea | 1.4 | 137.2 | 0.60 | 13.4 | 24.5 | 105.1 | 9.2 | 2.32 | 14.58 |
| Lebanon | 16.8 | 70.8 | 0.26 | 11.0 | 7.6 | 99.8 | 2.5 | 2.04 | 16.18 |
| Nepal | 1.2 | 120.8 | 0.35 | 13.0 | 19.6 | 112.0 | 2.6 | 2.18 | 16.89 |
| Netherlands | 1.7 | 117.4 | 0.60 | 12.3 | 19.8 | 93.9 | 6.1 | 2.62 | 17.56 |
| Poland | 3.5 | 73.6 | 0.30 | 8.2 | 9.1 | 90.0 | 2.7 | 2.39 | 15.63 |
| Romania | 2.8 | 104.1 | 0.50 | 9.6 | 18.4 | 93.0 | 6.0 | 2.55 | 15.53 |
| Russia | 2.3 | 94.3 | 0.38 | 9.8 | 19.8 | 89.2 | 4.9 | 2.41 | 16.23 |
| South Africa | 5.4 | 153.1 | 0.50 | 16.5 | 17.4 | 117.0 | 7.8 | 2.40 | 18.81 |
| USA | 4.3 | 128.5 | 0.52 | 15.0 | 15.9 | 114.0 | 6.1 | 2.45 | 15.78 |

(continued)

**Table 2.1** (continued)

| Province/ country | Till | Hgt (cm) | CulD (cm) | Node# | PanL (cm) | DaysM (day) | DWP (g) | DW1000 (g) | Prot (%) |
|---|---|---|---|---|---|---|---|---|---|
| Average | 3.3 | 111.9 | 0.4 | 11.8 | 17.6 | 101.4 | 5.2 | 2.4 | 16.3 |
| CV% | 101.39 | 23.06 | 27.72 | 23.55 | 29.56 | 10.83 | 50.08 | 8.69 | 8.17 |

Modified from Li et al. (1995)
Abbreviations: *Till* number of tillers including main culm, *Hgt* plant height to panicle base, *CulD* diameter of main culm base at maturity, *Node#* visible node number of main culm, *PanL* length of main panicle, *DaysM* days from emergence to maturity, *DWP* grain dry weight per plant including tillers, *DW1000* 1000-grain dry weight, *Prot* crude protein content, *CV%* coefficient of variation (=standard deviation divided by the mean)×100 %

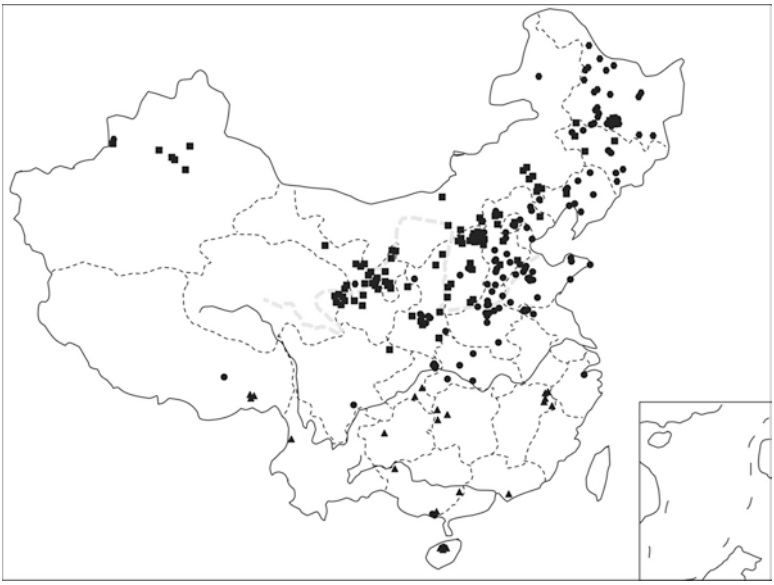

**Fig. 2.1** Geographical distributions of subpopulations of Chinese foxtail millet landraces. *Black triangle*—SCR Group (all provinces south of Qinling Mountain); *black circle*—SSSR Group (Henan, Hebei, Shandong, Beijing, Tianjin, and connected parts of Jiangsu, Liaoning, and Jilin provinces); *black square*—SR Group (Shanxi, Shannxi, Inner Mongolia, Gansu, Ningxia, Xianjiang, and western part of Jilin and Liaoning provinces); *black hexagon*—ESR Group (northeast Heilongjiang province). The Yellow River valley is marked a *grey* broken line (modified from Wang et al. 2012)

were identified, which suggest highly diversified genomic variants within and among Chinese ecotypes. Four subpopulations (Fig. 2.1) were inferred, corresponding to the early-spring sowing region (northeast Heilongjiang province, where foxtail millet is sowed in late April or early May annually), spring sowing region (Shanxi, Shannxi, Inner Mongolia, Gansu, Ningxia, Xianjiang, and western part of Jilin and Liaoning provinces, where foxtail millet is sowed in May), summer and spring sowing region (Henan, Hebei, Shandong, Beijing, Tianjin, and connected parts of Jiangsu, Liaoning

and Jilin provinces, where foxtail millet can be sowed from late April to early July), and southern china region (all provinces south of Qinling Mountain, where foxtail millet can be sowed even in late summer and early autumn). The highest level of molecular diversity was detected in accessions grouped into subpopulations sourced from eco-regions around the Yellow River valley in north China.

### 2.2.2 Elite Varieties

China is considered as the center of foxtail millet production and most crop improvement efforts. Since the 1950s, over 550 foxtail millet elite varieties developed by more than 40 breeding programs were released and registered with local and national authorities in China. According to studies by Jia et al. (2013a), phenotypic changes involved in modern breeding of foxtail millet include yield-related traits such as grain yield per plant, grain yield per main stem, panicle weight per main stem, panicle weight per plant, 1000-grain weight, and number of total panicles per plant. Using microsatellite markers, Jia et al. (2015) sampled 348 accessions and detected an average of 17.9 allele number per locus in Chinese elite varieties.

Two clear subgroups, designated as spring-sowing and summer-sowing types, were identified by Jia et al. (2015). These correspond to the eco-regions of the two main foxtail millet production areas in China. In addition, analyses of subgroups associated with breeding histories, breeding preference, and planting configuration indicated that the genetic background of most elite varieties released in recent years was from summer-sowing ecotypes, as a result of the response of breeders to the recent increase of the multi-cropping index (number of crops that can be sown and harvested in a year) in China (Xin et al. 2009).

The spring-sowing cluster is comprised of varieties developed prior to the 1980s, of which most are not currently used in grain production. The summer-sowing cluster is composed of elite varieties that were mostly developed after the 1980s. Elite varieties in the spring-sowing cluster possess higher levels of genetic diversity compared to those in the summer-sowing cluster, and varieties from the same breeding programs and similar eco-environmental conditions tend to be more closely related. Chinese foxtail millet elite varieties released by breeding programs conducted in Shanxi and Jilin provinces were dissimilar from most other lines, suggesting that unique germplasm was incorporated into these breeding programs (Jia et al. 2015).

### 2.2.3 Wild Relatives

Green foxtail (*Setaria viridis*), the ancestor and wild relative of cultivated foxtail millet, is distributed worldwide. To date, little morphological research on the diversity of Chinese green foxtail populations has been published. Investigations of genetic diversity of Chinese green foxtail accessions using transposon display (Hirano et al. 2011), intersimple sequence repeat (Li et al. 2012), and amplified

fragment length polymorphism markers (Le Thierry d'Ennequin et al. 2000) have found very high levels of molecular diversity. Green foxtail is also becoming a model for functional genomics studies, focusing on abiotic stress tolerance (Qie et al. 2014), crop domestication (Doust et al. 2014), and grass evolution (Doust et al. 2009). The well-studied green foxtail accession A10 that has been used as model for morphological development and $C_4$ photosynthesis research was collected in Canada, but is no doubt originally from China (Brutnell et al. 2010; Li and Brutnell 2011; Bennetzen et al. 2012; Caemmerer et al. 2012).

Genetic characterization of DNA sequence polymorphisms from green foxtail germplasm is pivotal for analyzing domestication, evolution, and potential for breeding in this wild grass species (Huang et al. 2014). In the analysis reported by Jia et al. (2013b), the microsatellite diversity of Chinese green foxtail accessions is 50 % higher than cultivated foxtail millet, with an average of 33.5 alleles per locus. Subclusters in the south and north are inferred according to eco-geographical regions designated in China, and higher molecular diversity was found in the northern subcluster. Phylogenetic studies of both Chinese foxtail millet and green foxtail samples also suggest that North China may be the first domestication center, or at least one of the oldest locations for cultivation of this ancient cereal crop. Genetic introgression analysis illustrates lower levels of gene flow within green foxtail subclusters coupling with higher frequency of introgressions between cultivated and wild Chinese Setaria accessions (Fig. 2.2). Similar results were also detected by Wang et al. (2010) through re-sequencing of nine genomic regions in both wild and cultivated Setaria samples. Decay of LD was found to be faster in *S. viridis* than *S. italica* (Wang et al. 2012), which is consistent with higher rates of cross pollination

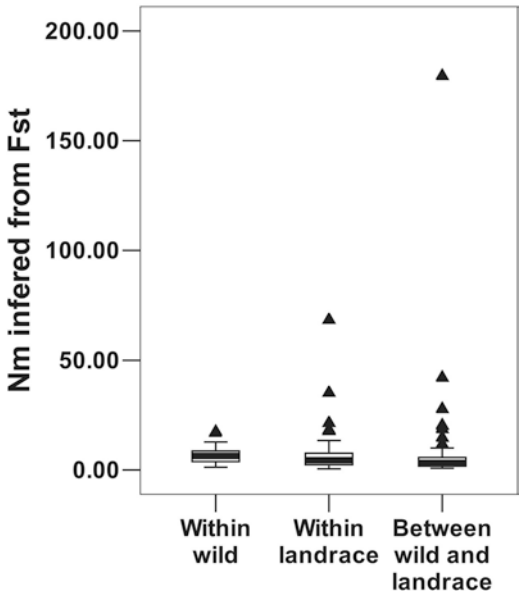


**Fig. 2.2** Gene flow identified by Nm inferred from classical *F* test within and between green foxtail and foxtail millet landraces (modified from Jia et al. 2013b)

within wild Setaria accessions. Using high density SNPs markers, Huang et al. (2014) also detected a decline in LD decay distance (<45 Kb) compared with that of cultivated Setaria accessions (~100 Kb) (Jia et al. 2013a).

Using simulation analysis (Wang et al. 2010), a low level of gene flow from cultivated landraces to wild relatives was detected, but the reverse process was even smaller. Gene flow between *S. italica* and *S. viridis* was also identified through microsatellite analysis (Jia et al. 2013b) and introgressions of herbicide-resistant genes detected in the field (Shi et al. 2008). These observations illustrate that under long-term intensive human selection, alleles transferred from wild green foxtail to cultivated foxtail millet were mostly eliminated, but genes introgressed into the wild relatives might be retained at a higher level, as evidenced by many weedy types (labeled as "giant foxtail") that are morphologically similar to cultivated foxtail millet in and around foxtail millet fields (Li et al. 1942, 1945).

## 2.3 Selection Bottlenecks

Domestication and improvement of ancient cereal crops always leads to a dramatic decline in morphological diversity and genetic polymorphisms. Characteristics like loss of shattering, less tillering, plant height variation, and grain yield improvement have been widely identified in crop species such as rice, wheat, maize, barley, etc. For Setaria, morphological changes from the wild green foxtail to cultivated landraces include larger panicles, higher grain yield, less tillering/branching, longer period of flowering, loss of seed shattering and dormancy, and variation in biotic/abiotic stress tolerance, etc. (Doust et al. 2004, 2005, 2014; Brutnell et al. 2010; Mauro-Herrera et al. 2013; Qie et al. 2014; Mauro-Herrera and Doust 2016).

Through re-sequencing of nine genomic regions of wild and cultivated Setaria, Wang et al. (2010) reported a 55 % loss of nucleotide diversity from wild relatives to cultivated Setaria accessions. To date, only one candidate gene (*Si037789m*) has been identified for the control of seed shattering in Setaria, based on comparative genetic analysis (Lin et al. 2012) and an 855-bp insertion or deletion in coding regions confirmed between *Setaria italica* and *Setaria viridis* (Jia et al. 2013a). Furthermore, several candidate loci involved in controlling branching (Doust et al. 2004; Mauro-Herrera and Doust 2016), panicle formation (Doust et al. 2005), and flowering time (Mauro-Herrera et al. 2013; Doust et al. 2014) have been identified in Setaria by QTL mapping analysis (see Chap. 12).

Marked changes have also been observed between Chinese traditional landraces and modern cultivars, especially in yield and morphological traits such as panicle weight, tiller number, hull color, and leaf angle. A total of 36 genomic regions corresponding to improvement of elite Setaria varieties have been characterized, but no genes have yet been identified (Jia et al. 2013a).

Genetic bottlenecks between wild relatives, landraces, and elite varieties have also been detected by microsatellite variation (Wang et al. 2012; Jia et al. 2013b, 2015). Studies using the same set of SSR markers identified an average of 33.5

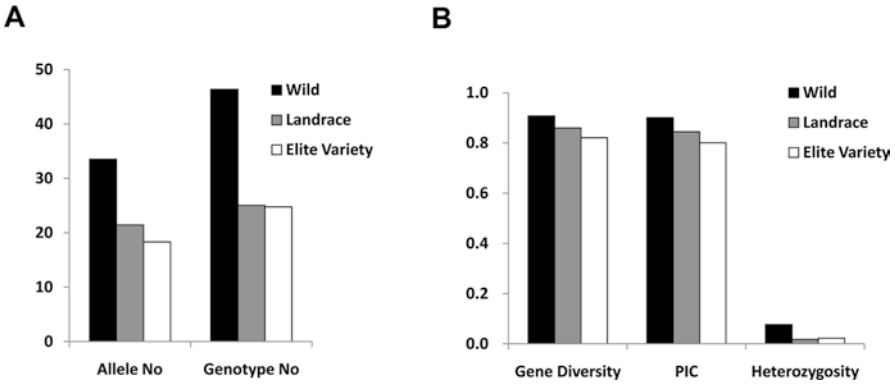


**Fig. 2.3** Selective bottlenecks detected in Setaria. Allele and genotype no. (**a**) and gene diversity, PIC values, and heterozygosity (**b**) per locus detected in three panels of Setaria gene pools (modified from Jia et al. 2015)

alleles per locus in Chinese green foxtail, 20.9 alleles per locus in Chinese foxtail millet landraces, and 17.8 alleles per locus in Chinese elite foxtail millet varieties. This result is consistent with a loss of genetic diversity between landrace and elite gene pools as a result of two levels of genetic bottlenecks during the development of modern foxtail millet elite varieties in China. Selection for the domestication of green foxtail appears to be stronger than that for the improvement of cultivated foxtail millet (Fig. 2.3). Genomic regions contributing to the domestication process of Setaria have also received attention in recent years (Doust et al. 2014). For elite varieties, selective sweep regions of foxtail millet improvement are ~200 Kb in average, and selective pressures are modest (the largest $\pi_l/\pi_m$ value is 6), comparable to what has been observed in maize (*Zea mays*) (Jia et al. 2013a).

## 2.4 QTLs Revealed by GWAS

Quantitative trait loci (QTLs) are genomic regions containing or linked to genes controlling or involved in a quantitative trait. Mapping of QTLs is an important first step for functional gene isolation, metabolic pathway construction, and marker-assisted selection (MAS), as well as pyramiding of target phenotypes in breeding programs. In foxtail millet, recombinant segregating population-based linkage analysis (Doust et al. 2004, 2005, 2014; Mauro-Herrera et al. 2013; Qie et al. 2014; Mauro-Herrera and Doust 2016) and natural accession-based association analysis (Jia et al. 2013a, 2015; Gupta et al. 2014) have been used to localize QTLs into specific genomic regions that are involved with phenotypes. Over 900 QTLs (Table 2.2) controlling 47 agronomical important characters have been identified under diverse environmental conditions by GWAS, using over a million SNPs and microsatellite markers (Jia et al. 2013a, 2015; Gupta et al. 2014).

**Table 2.2** Selected results of GWAS for QTL detection in Setaria

| Trait | Chr. | Marker name/physical position | Location of trials | Reference |
|---|---|---|---|---|
| Anther color | 6 | 34,378,428 | Chaoyang, Beijing, Changzhi and Anyang of China | Jia et al. (2013a) |
| Branch number per main stem | 2 | 2,465,005 | Chaoyang, China | Jia et al. (2013a) |
| | 5 | 10,414,336 | Beijing, China | Jia et al. (2013a) |
| | 9 | 44,633,850 | Anyang, China | Jia et al. (2013a) |
| | 1 | 27,146,578 | Changzhi, China | Jia et al. (2013a) |
| | 5 | 31,851,805 | Sanya, China | Jia et al. (2013a) |
| Tiller number | 8 | 35,517,380 | Chaoyang, China | Jia et al. (2013a) |
| | 2 | 2,364,917 | Beijng, China | Jia et al. (2013a) |
| | 2 | 2,374,147 | Anyang, China | Jia et al. (2013a) |
| | 9 | 56,850,515 | Changzhi, China | Jia et al. (2013a) |
| | 2 | 2,411,360 | Changzhi, China | Jia et al. (2013a) |
| 1000-grain weight | 1 | 37,243,809 | Anyang, China | Jia et al. (2013a) |
| | 1 | b260/37,750,853~37,751,008 | New Delhi, India | Gupta et al. (2014) |
| | 5 | P75/— | New Delhi, India | Gupta et al. (2014) |
| | 5 | b129/7,959,308~7,959,430 | New Delhi, India | Gupta et al. (2014) |
| Bristle length | 1 | 37,748,213 | Beijing, Anyang, Changzhi of China | Jia et al. (2013a) |
| | 1 | 37,751,602 | Sanya, China | Jia et al. (2013a) |
| Heading date | 2 | 49,164,136 | Beijing and Anyang of China | Jia et al. (2013a) |
| | 1 | 37,215,226 | Anyang, China | Jia et al. (2013a) |
| | 8 | 28,507,351 | Changzhi, China | Jia et al. (2013a) |
| | 2 | 2,711,177 | Sanya, China | Jia et al. (2013a) |
| | 9 | P44/— | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| Leaf length | 2 | 49,119,122 | Beijing, China | Jia et al. (2013a) |
| | 2 | 49,117,404 | Anyang, China | Jia et al. (2013a) |

(continued)

**Table 2.2** (continued)

| Trait | Chr. | Marker name/physical position | Location of trials | Reference |
|---|---|---|---|---|
| Panicle length | 2 | 22,845,341 | Beijing, China | Jia et al. (2013a) |
| | 1 | 37,343,439 | Anyang, China | Jia et al. (2013a) |
| | 4 | b109/1,290,206~1,290,107 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 1 | b182/32,831,786~32,831,986 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 9 | b217/— | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 4 | b236/37,687,774~37,687,917 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 4 | b247/— | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 4 | b266/40,054,016~40,053,821 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 9 | b269/— | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 1 | P8/37,804,404~37,804,237 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 5 | b129/7,959,308~7,959,430 | New Delhi, India | Gupta et al. (2014) |
| Bristle color | 4 | 7,400,384 | Chaoyang, China | Jia et al. (2013a) |
| | 4 | 7,390,482 | Beijing, China | Jia et al. (2013a) |
| | 4 | 7,390,404 | Anyang, China | Jia et al. (2013a) |
| | 4 | 7,215,041 | Changzhi, China | Jia et al. (2013a) |
| | 4 | 7,284,766 | Sanya, China | Jia et al. (2013a) |
| Leaf sheath color | 7 | 26,883,452 | Chaoyang, China | Jia et al. (2013a) |
| | 7 | 26,908,254 | Beijing, China | Jia et al. (2013a) |
| | 4 | 7,322,156 | Beijing, China | Jia et al. (2013a) |
| | 7 | 26,859,601 | Anyang, China | Jia et al. (2013a) |
| | 4 | 7,349,812 | Anyang, China | Jia et al. (2013a) |
| | 7 | 26,910,276 | Changzhi, China | Jia et al. (2013a) |
| | 4 | 7,390,404 | Changzhi, China | Jia et al. (2013a) |
| | 7 | 26,908,030 | Sanya, China | Jia et al. (2013a) |
| | 4 | 7,322,156 | Sanya, China | Jia et al. (2013a) |

(continued)

**Table 2.2** (continued)

| Trait | Chr. | Marker name/physical position | Location of trials | Reference |
|---|---|---|---|---|
| Hull color | 1 | 5,480,917 | Chaoyang, China | Jia et al. (2013a) |
| | 9 | 54,490,294 | Chaoyang, China | Jia et al. (2013a) |
| | 1 | 5,467,847 | Anyang, China | Jia et al. (2013a) |
| | 9 | 54,533,849 | Anyang, China | Jia et al. (2013a) |
| | 1 | 5,476,983 | Changzhi, China | Jia et al. (2013a) |
| | 9 | 54,490,656 | Changzhi, China | Jia et al. (2013a) |
| Height of main stem (plant height) | 5 | MPGA31/27,433,840~27,433,863 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 9 | X4/98,538~98,850 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| Peduncle length | 4 | P2/1,221,537~1,221,682 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 3 | b225/11,756,544~11,756,666 | New Delhi, India | Gupta et al. (2014) |
| | 5 | b129/7,959,308~7,959,430 | New Delhi, India | Gupta et al. (2014) |
| | 5 | P75/— | New Delhi, India | Gupta et al. (2014) |
| Panicle diameters | 7 | b123/— | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| | 3 | P98/15,515,665~15,515,807 | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| Stem diameters | 7 | P29/— | Beijing, Changzhi and Anyang of China | Jia et al. (2015) |
| Early blooming date | 2 | 49,164,136 | Beijing, China | Jia et al. (2013a) |
| | 1 | 37,210,883 | Anyang, China | Jia et al. (2013a) |
| | 6 | SiGMS9645/— | New Delhi, India | Gupta et al. (2014) |
| Flag leaf width | 5 | b129/7,959,308~7,959,430 | New Delhi, India | Gupta et al. (2014) |
| | 7 | P59/— | New Delhi, India | Gupta et al. (2014) |
| | 8 | b185/— | New Delhi, India | Gupta et al. (2014) |
| Grain yield | 5 | P75/— | New Delhi, India | Gupta et al. (2014) |
| | 5 | b129/7,959,308~7,959,430 | New Delhi, India | Gupta et al. (2014) |
| Inflorescence bristles | 5 | b129/7,959,308~7,959,430 | New Delhi, India | Gupta et al. (2014) |
| Grain shape | 3 | P61/— | New Delhi, India | Gupta et al. (2014) |
| Inflorescence compactness | 3 | b225/11,756,544~11,756,666 | New Delhi, India | Gupta et al. (2014) |

Physical positions of QTLs loci were acquired according to reference genome sequence of Yugu1 (*Setaria italica* version 2.1) from website: http://phytozome.jgi.doe.gov/pz/portal.html

Results of QTL mapping suggest that several coloration-related traits are controlled by Mendelian genes, but that other quantitative traits including yield components, disease resistance, growth time, and morphological development are controlled by a number of genomic regions with diverse minor effects. Association mapping analysis also suggests that most marker and phenotype associations were environment-specific and only acted on a single trait (Jia et al. 2015). Venn plot analysis of QTLs identified under different environmental conditions (Jia et al. 2013a) implied that a majority of morphological characters were closely related to locally adapted responses in Setaria to various climatic and ecological environments, so that Setaria can potentially be used as the complementary model to *Arabidopsis* for future ecological genomics studies.

## 2.5 Perspectives

Domestication and breeding processes have profoundly influenced the genetic diversity and population structure present in Setaria. Understanding the genetic basis of domestication and phenotypic variation in Setaria can help us efficiently utilize these genetic resources for variety improvement and relevant biological studies including $C_4$ photosynthesis, stress tolerance, plant architecture, and panicle development.

Morphological and genetic diversity detected in Setaria germplasm suggests that the sequence diversity ($\pi$) of cultivated Setaria accessions is approximately ~0.0010, on a similar scale to that found in rice. A high level of diversity and the ability of Setaria to adapt to extreme local environments will make Setaria an attractive model for agronomic studies of crop species like maize (*Zea mays*), wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), sorghum (*Sorghum bicolor*), switchgrass (*Panicum virgatum*), and oat (*Avena sativa*).

Genome variations in Setaria can be characterized by genotyping each individual accession. Recently, many kinds of molecular markers have been developed in Setaria (Jia et al. 2009; Yadav et al. 2015; B et al. 2013; Zhang et al. 2014; Pandey et al. 2013; Muthamilarasan et al. 2014; Gupta et al. 2012; Wang et al. 1998; Gale and Devos 1998). New generations of genotyping technologies (Sequencing-based genotyping, genotyping by sequencing, RNA-seq-based genotyping, Exon-sequencing-based genotyping, etc.) have also been developed in recent years (Huang and Han 2014), including a low coverage whole-genome re-sequencing method for analyzing the origin and domestication of rice (Huang et al. 2012) and construction of haplotype maps in rice (Huang et al. 2010) and Setaria (Jia et al. 2013a). We believe that these genomic approaches will continue to reveal variability in the genetic architecture of valuable traits in Setaria in the future. Likewise, with continued significant decreases in sequencing costs, we can predict that more and more such sequencing data sets will be available to facilitate GWAS and domestication studies in Setaria.

## References

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye C, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30:556–61.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X-G, Kellogg E, Eck JV. *Setaria viridis*: a model for C4 photosynthesis. Plant Cell. 2010;22:2537–44.

Caemmerer SV, Quick WP, Furbank RT. The development of $C_4$ rice: current progress and future challenges. Science. 2012;336:1671–2.

Cao NG. Selection and breeding of cereals, fruit trees and husbandry animals in ancient China. In: Guo WT, Cao NG, Song ZQ, Ma XQ, editors. Traditional agriculture and modern agriculture in China. Beijing: Agricultural Scientech Publishing House; 1986. p. 169–201.

Diao X, Schnable J, Bennetzen JL, Li J. Initiation of *Setaria* as a model plant. Front Agric Sci Eng. 2014;1(1):16–20.

Doust AN, Devos KM, Gadberry M, Gale MD, Kellogg EA. Genetic control of branching in the foxtail millet. Proc Natl Acad Sci U S A. 2004;101:9045–50.

Doust AN, Devos KM, Gadberry M, Gale MD, Kellogg EA. The genetic basis for inflorescence variation between foxtail and green millet (*Poaceae*). Genetics. 2005;169:1659–72.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL. Foxtail millet: a sequence-driven grass model system. Plant Physiol. 2009;149:137–41.

Doust AN, Lukens L, Olsen KM, Mauro-Herrera M, Meyer A, Rogers K. Beyond the single gene: how epistasis and gene-by-environment effects influence crop domestication. Proc Natl Acad Sci U S A. 2014;111:6178–83.

Gale MD, Devos KM. Comparative genetics in the grasses. Proc Natl Acad Sci U S A. 1998;95: 1971–4.

Gupta S, Kumari K, Sahu PP, Vidapu S, Prasad M. Sequence-based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* (L.) Beauv]. Plant Cell Rep. 2012;31:323–37.

Gupta S, Kumari K, Muthamilarasan M, Parida SK, Prasad M. Population structure and association mapping of yield contributing agronomic traits in foxtail millet. Plant Cell Rep. 2014;33:881–93.

Hirano R, Naito K, Fukunaga K, Watanabe KN, Ohsawa R, Kawase M. Genetic structure of landraces in foxtail millet (*Setaria italica* (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet. Genome. 2011;54:498–506.

Huang X, Han B. Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol. 2014;65:531–51.

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang Q, Li J, Han B. Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet. 2010;42:961–7.

Huang X, Kurata N, Wei X, Wang Z, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, Guo Y, Lu Y, Zhou C, Fan D, Weng Q, Zhu C, Huang T, Zhang L, Wang Y, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan X, Xu Q, Dong G, Zhan Q, Li C, Fujiyama A, Toyoda A, Lu T, Feng Q, Qian Q, Li J, Han B. A map of rice genome variation reveals the origin of cultivated rice. Nature. 2012;490:497–501.

Huang P, Feldman M, Schroder S, Bahri B, Diao X, Zhi H, Estep M, Baxter I, Devos KM, Kellogg EA. Population genetics of *Setaria viridis*, a new model system. Mol Ecol. 2014;23:4912–25.

Jia X, Zhang Z, Liu Y, Zhang C, Shi Y, Song Y, Wang T. Development and genetic mapping of SSR markers in foxtail millet [*Setaria italica* (L.) P. Beauv.]. Theor Appl Genet. 2009;118:821–9.

Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, Chai Y, Yang L, Liu K, Lu H, Zhu C, Lu Y, Zhou C, Fan D, Weng Q, Guo Y, Huang T, Zhang L, Lu T, Feng Q, Hao H, Liu H, Lu P, Zhang N, Li Y, Guo E, Wang S, Wang S, Liu J, Zhang W, Chen G, Zhang B, Li W, Wang Y, Li H, Zhao B, Li J, Diao X, Han B. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013a;45:957–61.

Jia G, Shi S, Wang C, Niu Z, Chai Y, Zhi H, Diao X. Molecular diversity and population structure of Chinese green foxtail [*Setaria viridis* (L.) Beauv.] revealed by microsatellite analysis. J Exp Bot. 2013b;64:3645–55.

Jia G, Liu X, James CS, Niu Z, Wang C, Li Y, Wang S, Wang S, Liu J, Guo E, Zhi H, Diao X. Microsatellite variations of elite Setaria varieties released during last six decades in China. PLoS One. 2015;10(5):e0125688.

Jusuf M, Pernes J. Genetic variability of foxtail millet (*Setaria italica* P. Beauv.). Theor Appl Genet. 1985;63:117–9.

Le Thierry d'Ennequin M, Panaud O, Toupance B, Sarr A. Assessment of genetic relationships between *Setaria italica* and its wild relative *S. viridis* using AFLP markers. Theor Appl Genet. 2000;100:1061–6.

Li P, Brutnell TP. *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. J Exp Bot. 2011;62:3031–7.

Li Y, Wu S. Traditional maintenance and multiplication of foxtail millet (*Setaria italica* (L.) P. Beauv.) landraces in China. Euphytica. 1996;87:33–8.

Li C, Pao W, Li H. Interspecific crosses in *Setaria*. J Hered. 1942;33:351–5.

Li H, Li C, Pao W. Cytogenetical and genetical studies of the interspecific cross between the cultivated foxtail millet, *Setaira italica* (L.) Beauv. and the green foxtail millet *S. viridis* L. J Am Soc Agron. 1945;37:32–54.

Li Y, Wu S, Cao Y. Cluster analysis of an international collection of foxtail millet (*Setaria italica* (L.) P. Beauv.). Euphytica. 1995;83:79–85.

Li Y, Jia J, Wang Y, Wu S. Intraspecific and interspecific variation in *Setaria* revealed by RAPD analysis. Genet Resour Crop Evol. 1998;45:279–85.

Li W, Zhi H, Wang Y, Li H, Diao X. Assessment of genetic relationship of foxtail millet with its wild ancestor and close relatives by ISSR markers. J Integr Agric. 2012;11(4):556–66.

Lin Z, Li X, Shannon LM, Yeh C, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J. Parallel domestication of the *Shattering1* genes in cereals. Nat Genet. 2012;44:720–4.

Lu H, Yang X, Ye M, Liu K, Xia Z, Ren X, Cai L, Wu N, Liu T. Culinary archaeology: millet noodles in late Neolithic China. Nature. 2005;437:967–8.

Lu H, Zhang J, Liu KB, Wu N, Li Y, Zhou K, Ye M, Zhang T, Zhang H, Yang X, Shen L, Xu D, Li Q. Earliest domestication of common millet (*Panicum miliaceum*) in east Asia extended to 10,000 years ago. Proc Natl Acad Sci U S A. 2009;106:7367–72.

Mauro-Herrera M, Doust AN. Development and genetic control of plant architecture and biomass in the panicoid grass, *Setaria*. PLoS One. 2016;11(3):e0151346.

Mauro-Herrera M, Wang X, Barbier H, Brutnell TP, Devos KM, Doust AN. Genetic control and comparative genomic analysis of flowering tine in *Setaria* (*Poaceae*). G3 Genes Genom Genet. 2013;3:283–95.

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc Natl Acad Sci U S A. 2013;110:453–8.

Muthamilarasan M, Venkata SB, Pandey G, Kumari K, Parida SK, Prasad M. Development of 5123 intron-length polymorphic markers for large-scale genotyping application in foxtail millet. DNA Res. 2014;21:41–52.

Pandey G, Misra G, Kumari K, Gupta S, Parida SK, Chattopadhyay D, Prasad M. Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. DNA Res. 2013;20:197–207.

Qie L, Jia G, Zhang W, Schnable J, Shang Z, Li W, Liu B, Li M, Chai Y, Zhi H, Diao X. Mapping of quantitative trait locus (QTLs) that contribute to germination and early seedling drought tolerance in the interspecific cross *Setaria italica* × *Setaria viridis*. PLoS One. 2014;9(7):e101868.

Reddy VG, Upadhyaya HD, Gowda CLL. Characterization of world's foxtail millet germplasm collections for morphological traits. J SAT Agric Res. 2006;2:1–3.

Schontz D, Rether B. Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv.: identification and classification of lines with RAPD markers. Plant Breed. 1999;118:190–2.

Shi Y, Wang T, Li Y, Darmency H. Impact of transgene inheritance on the mitigation of gene flow between crops and their wild relatives: the example of foxtail millet. Genetics. 2008;180:969–75.

B VS, Muthamilarasan M, Misra G, Prasad M. FmMDb: a versatile database of foxtail millet markers for millets and bioenergy grasses research. PLoS One. 2013;8:e71418.

Van K, Onoda S, Kim MY, Kim KD, Lee SH. Allelic variation of the Waxy gene in foxtail millet [*Setaria italica* (L.) P. Beauv.] by single nucleotide polymorphisms. Mol Genet Genomics. 2008;279:255–66.

Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD. Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1998;96:31–6.

Wang C, Chen J, Zhi H, Yang L, Li W, Wang Y, Li H, Zhao B, Chen M, Diao X. Population genetics of foxtail millet and its wild ancestor. BMC Genet. 2010;11:90.

Wang C, Jia G, Zhi H, Niu Z, Chai Y, Li W, Wang Y, Li H, Lu P, Zhao B, Diao X. Genetic diversity and population structure of Chinese foxtail millet [*Setaria italica* (L.) Beauv.] landraces. G3 Genes Genom Genet. 2012;2:769–77.

Xin L, Li X, Zhu H, Tan M. China's potential of grain production due to changes in agricultural land utilization in recent years. Chin Geogr Sci. 2009;19(2):97–103.

Yadav CB, Bonthala VS, Muthamilarasan M, Pandey G, Khan Y, Prasad M. Genome-wide development of transposable elements-based markers in foxtail millet and construction of an integrated database. DNA Res. 2015;22:79–90.

Yang X, Wan Z, Perry L, Lu H, Wang Q, Zhao C, Li J, Xie F, Yu J, Cui T, Wang T, Li M, Ge Q. Early millet use in northern china. Proc Natl Acad Sci U S A. 2012;109:3726–30.

Zhang S, Tang C, Zhao Q, Li J, Yang L, Qie L, Fan X, Li L, Zhang N, Zhao M, Liu X, Chai Y, Zhang X, Wang H, Li Y, Li W, Zhi H, Jia G, Diao X. Development of highly polymorphic simple sequence repeat markers using genome-wide microsatellite variant analysis in foxtail millet [*Setaria italica* (L.) P. Beauv]. BMC Genomics. 2014;15:78.

# Chapter 3
# Genetic Diversity and Geographic Distribution of North American *Setaria viridis* Populations

**Pu Huang and Maximillian Feldman**

**Abstract** *Setaria viridis* is an emerging model system that is used to study $C_4$ photosynthesis and many other traits relating to agricultural productivity of panicoid grain crops and bioenergy grasses. Originally from China, naturalized populations of *S. viridis* are widely distributed in North America and contain a high degree of genetic and phenotypic diversity. In this chapter, we estimate the distribution range of *S. viridis* in North America using a climate-based species distribution model. We then summarize some recent advances in understanding the genetic diversity, population structure, linkage disequilibrium, and biogeography of North American *S. viridis* populations. We further expound upon the necessity of expanding germplasm collections of *S. viridis,* our own ongoing collection efforts, and some of the sequence and germplasm resources currently being generated. Finally, we discuss utilizing these tools as a means for genetic dissection of complex traits, potential opportunities for gene discovery, and how our understanding of the natural variation in *S. viridis* populations will make agriculture a more productive and sustainable industry.

**Keywords** *Setaria viridis* • North America • Genetic variation • Species distribution modeling • Green foxtail

## 3.1 *Setaria viridis*, From a Weed to a Genetic Model

Green foxtail (*Setaria viridis*) is widely recognized as the direct wild ancestor and a crucial germplasm pool for the agricultural crop species foxtail millet (*S. italica*) in Asia and Europe (Barton et al. 2009; Wang et al. 2010). In North America, where foxtail millet is not a major crop, *S. viridis* (among a few other *Setaria* species) is most commonly recognized as a major invasive weed in corn and soybean fields (Wang et al. 1995; Délye et al. 2002; Dekker 2003). Accordingly, understanding the ecology and genetics of weed control traits in *S. viridis* has been a major focus of scientific inquiry

P. Huang (✉) • M. Feldman
Donald Danforth Plant Science Center, 975 North Warson Rd., St. Louis, MO 63132, USA
e-mail: phuang@danforthcenter.org



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_3

in recent years. Several studies have described the geographic origin of the weedy *S. viridis* populations in North America, including weed adaptation-related life history traits that enable rapid colonization of new territories and herbicide resistance (Dekker 2003). Among them, the genetic basis of herbicide resistance in *S. viridis* has been studied in most detail and candidate genes such as acetyl-CoA carboxylase have been identified (Délye et al. 2002, Darmency et al. chapter 15).

More recently, there has been a major shift in the focus on *S. viridis* in the scientific literature. The success and limitations of *Arabidopsis thaliana* as a model plant have led to the development of *Setaria viridis* as a model system for $C_4$ panicoid food crops (e.g. maize, sorghum, and sugarcane) and bioenergy grasses (e.g. switchgrass and *Miscanthus*) (Bennetzen et al. 2012; Brutnell et al. 2010, 2015; Doust et al. 2009). This new and exciting development has led to two important shifts in *S. viridis* research. First, the focus of interest has expanded from weed-related traits to include traits that determine agricultural productivity (e.g. photosynthetic efficiency, plant size/architecture, physiology, and abiotic stress tolerance). A majority of these traits exhibit a more complex genetic architecture than herbicide resistance (Brutnell et al. 2015; Doust et al. 2004). Second, detailed genetic dissection of traits in *S. viridis* is becoming increasingly more approachable with NextGen sequencing technologies. Thorough characterization of the genetic, physiological, and developmental mechanisms of trait variation in *S. viridis* can help define the potential trait value space, while additionally providing new leads for the manipulation of a trait. Understanding how such parameters influence agronomic traits in *S. viridis* will greatly improve the probability of successful translation of applications into economically important food and bioenergy crops (Brutnell et al. 2010). The success of these research efforts relies heavily on the development of new genetic tools and resources in this emerging model system. From the beginning of this decade, a wave of resource development has been building momentum. Techniques include standardized methods to perform genetic crosses, plant transformation, and development of RIL populations, as discussed in Chaps 10–12 and 19–21 of this book (Bennetzen et al. 2012; Brutnell et al. 2010, 2015; Huang et al. 2014; Jiang et al. 2013; Sebastian et al. 2014). The focus of this chapter, distribution and variation of natural populations of *S. viridis* in North America, is one component among these resources that is currently in a stage of rapid development. For simplicity, we refer to all naturalized, self-sustaining, and locally adapted *S. viridis* populations as "natural populations", regardless of their status as native, introduced, or invasive. They are distinguished from "mutant populations" which are created by artificial mutagenesis.

## 3.2 Natural Populations as a Source of Deep Genetic Variation

There are many salient reasons to study natural populations. Perhaps the most important is the vast amount of phenotypic and genetic variation that is maintained within these populations. While classical genetic approaches aim to

generate mutants with desired phenotypes or genotypes to study, evolution within natural populations provides massive amounts of standing variation. This is especially true for a widely distributed species like *S. viridis* (Figs. 3.1 and 3.2). As one of the most successful invasive plant species on Earth (Dekker 2003), *S. viridis* inhabits locations exhibiting large variance within multiple environmental variables (annual precipitation, temperature range, soil reactivity, etc.; Fig. 3.3). Based on this fact and results from studies performed with other plant species (Fournier-Level et al. 2011; Hancock et al. 2011; Long et al. 2013), it is tempting to speculate that colonization of such diverse environments is the result of genetic adaptation. Studying the mechanisms of adaptation within local populations will greatly improve our understanding of the solutions evolution has converged upon in response to environmental challenges. This knowledge may potentially translate to improvements in productivity and/or sustainability of related agricultural crops. For example, it is probably more likely to find a drought-resistant accession in regions with low annual precipitation than in regions with a wetter climate. Examining signatures of local adaptation to these environmental factors may identify genes contributing to these traits of interest (Fournier-Level et al. 2011; Hancock et al. 2011; Bergelson and Roux 2010; Evans et al. 2014; Wilson et al. 2015).

At the molecular genetic level, it is apparent that vast amounts of sequence variation persist in natural populations of *S. viridis* (Huang et al. 2014; Jia et al. 2013a). One advantage of exploring natural populations is that a small number of individuals capture a large number of polymorphisms relative to the number found in chemical mutagenesis populations. For example, in a panel of 144 diverse *S. viridis* accessions, there are on average 1200 nonsynonymous mutations per accession compared to the reference accession A10.1, while in a mutagenesis population of the same size there are less than 150 nonsynonymous mutations per mutant line (Huang et al. unpublished data). Accordingly, natural variation offers a powerful complement to mutagenesis, where the number of polymorphisms scales approximately linearly with the population size (Jiang et al. personal communication).

In addition, high levels of genetic diversity not only provide potential for variable phenotypes, but also densely spaced genetic markers to effectively detect recombination events. This is an important prerequisite to genome-wide association studies (GWAS). GWAS test the association between existing genetic polymorphisms and phenotypic variation in natural populations to identify genomic regions (QTLs) that are responsible for phenotypic variance within the population. Compared to the traditional linkage mapping (QTL mapping) approach, existing historical recombination in natural populations allows GWAS to achieve a much higher mapping resolution, sometimes down to single gene level (Fournier-Level et al. 2011; Evans et al. 2014; Jia et al. 2013b). The generation of high-density genetic markers is becoming increasingly feasible due to development of new low-cost sequencing technologies. As such, GWAS is becoming a widely used approach to resolve phenotype–genotype relationships and guide gene discovery. Understanding the existing genetic variation in natural

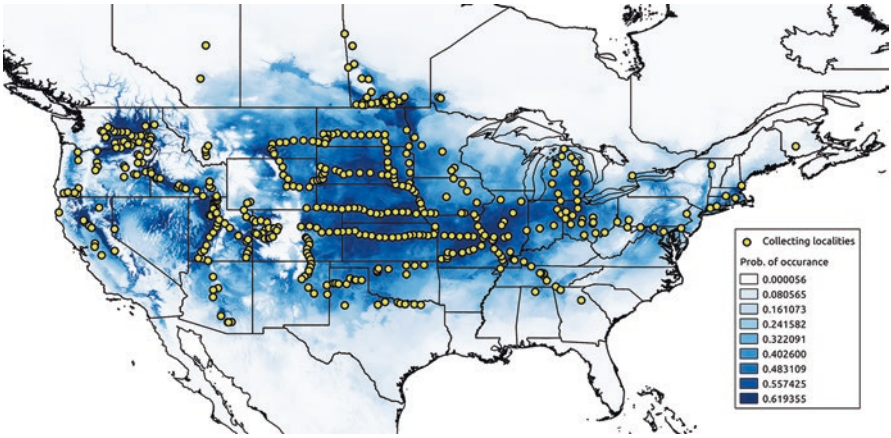


**Fig. 3.1** Collecting locality and species distribution modeling of North American *S. viridis*. *Open circles* indicate known collecting localities of *S. viridis* accessions. The *shadow regions* show the probable distribution range of *S. viridis*. *Darker color* corresponds to higher probability of occurrence

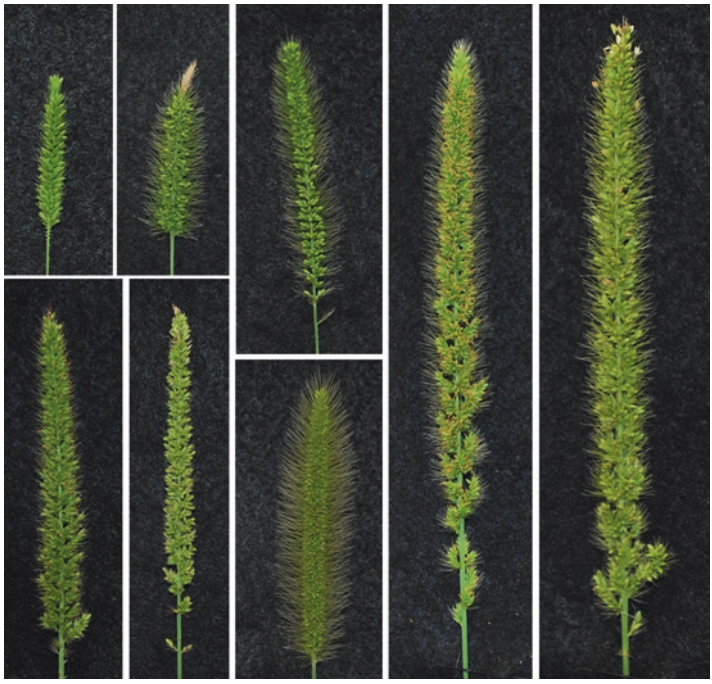

**Fig. 3.2** Morphological diversity in panicle shape and color

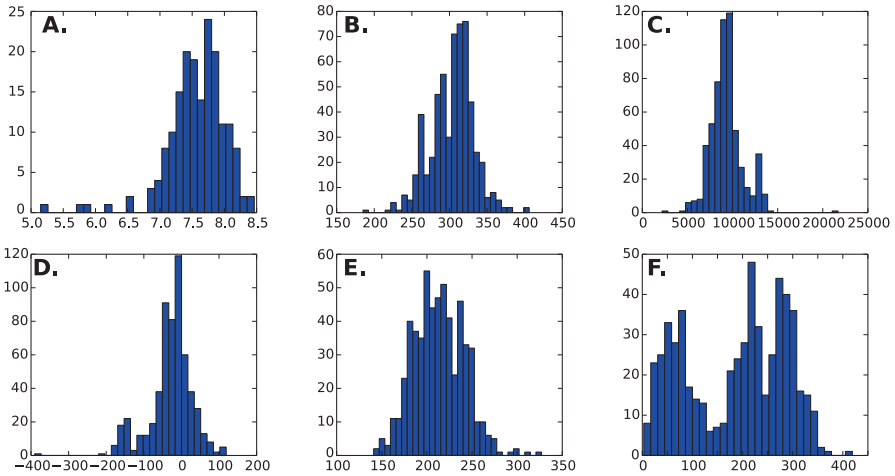


**Fig. 3.3** Distribution of soil pH and five climatic variables at collection sites. (**a**) soil pH, (**b**) maximum temperature of warmest month, (**c**) temperature seasonality, (**d**) mean temperature of coldest quarter, (**e**) mean temperature of warmest quarter, (**f**) precipitation of warmest quarter

populations is also important to control for factors leading to false positive hits in GWAS, such as population structure and long range linkage disequilibrium (Brachi et al. 2011).

## 3.3 A Species Distribution Model for North American *S. viridis*

One of the first steps to understand natural variation within *S. viridis* is to define its distribution range. This information provides guidelines for allocating resources during field collection efforts and may identify environmental trends that influence species habitat preference. Many herbarium records and on-line databases are available for *S. viridis* that can help guide further collection efforts (e.g. USDA http://plants.usda.gov/; GBIF http://www.gbif.org/; TROPICOS http://www.tropicos.org/). These databases provide crucial basic information about where to find *S. viridis*. However, due to a few inevitable problems these database resources should be used with caution. One of the major potential pitfalls is that specimens are sometimes mistakenly identified, due to the high resemblance between *S. viridis* and closely related *Setaria* species, such as *S. faberi* and *S. pumila*. Secondly, estimations of species distribution range may prove inaccurate due to uneven spatial distribution of collection efforts (e.g. false non-occurrence due to under sampling). Finally, in some cases distribution maps are not precisely geo-referenced with longitude/latitude coordinates, and thus may not be detailed enough to guide collections or infer climate characteristics.

Since proposing *S. viridis* as a model system, several groups have contributed germplasm to build a diversity panel of *S. viridis* ecotypes spanning North America (Huang et al. 2014; Huang et al. unpublished data). The species identities of most of these accessions have been verified through detailed examination of their morphology as well as through molecular genetic sequencing techniques. Summarizing the collections from our own collection efforts, as well as those kindly shared by various collaborators, a detailed distribution map of *S. viridis* in North America (Fig. 3.1) can be generated using a species distribution model (SDM) approach (Elith and Leathwick 2009; Phillips and Dudík 2008). This distribution map provides a quantitative measure of the probability of occurrence for North American *S. viridis*. From this distribution map, there are a few general conclusions. First, *S. viridis* is widely distributed in North America, but its distribution is not uniform. The central plains area of North America hosts the largest continuous *S. viridis* distribution. Second, *S. viridis* tends not to occur in high altitude regions (above ~2500 m), particularly within high altitude regions of the western US (Montana, Idaho, Wyoming, and Colorado; Fig. 3.1). Third, North American *S. viridis* is not found frequently in the southern US regions below latitude of 30°N (discussed below). There also seems to be a northern boundary of the species distribution in Southern Canada. However, the actual species distribution range potentially extends further north, as the collection effort in Canada remains localized and areas outside this region are likely under sampled.

SDMs have been widely used to predict the distribution range of particular species of interest. In this analysis, we used 19 bioclimatic variables from worldclim ((http://www.worldclim.org/); (Hijmans et al. 2005)) and package MaxEnt (Elith and Leathwick 2009; Phillips and Dudík 2008) to build SDMs. The relative importance of different variables contributing to the SDM was also evaluated by MaxEnt. Using two different assessment criteria (percent contribution and permutation importance; Table 3.1), variable Bio_5, or maximum temperature of warmest month contributed most to the SDM of North American *S. viridis*. This indicates that heat could be a primary factor limiting *S. viridis* distribution in the Southern US. This also provides an interesting contrast to distributions of *S. viridis* from other parts of the world. Notably, in China, *S. viridis* seems to be distributed to the more southern latitudes (Jia et al. 2013a), indicating that the local populations perhaps are adapted to the more subtropical climatic conditions there.

The genetic basis of climatic adaptation has been a focus of several recent studies (Fournier-Level et al. 2011; Hancock et al. 2011; Evans et al. 2014). A number of studies have also discussed the genetic basis of climatic adaptation in grasses, such as maize, sorghum, and rice (Gore et al. 2009; Huang et al. 2010; Morris et al. 2013). However, most of these species are domesticated crops. As such, climatic adaptation could be confounded by human domestication efforts. From this perspective, *S. viridis*, given its wide distribution range and increasing genomic resources, may provide an ideal opportunity to examine climatic adaptations in a $C_4$ Panicoid grass. This knowledge in turn could be further translated into food or energy crops, making them more sustainable and locally adapted.

**Table 3.1** Bioclimatic variables used to build species distribution model and their relative importance

| Variable | Interpretation[a] | Percent contribution[b] | Permutation importance[b] |
|---|---|---|---|
| Bio_05 | Max temperature of warmest month | 50.9 | 32 |
| Bio_11 | Mean temperature of coldest quarter | 11.1 | 6.5 |
| Bio_10 | Mean temperature of warmest quarter | 10.3 | 6.8 |
| Bio_01 | Annual mean temperature | 7.9 | 5.2 |
| Bio_03 | Isothermality | 3.4 | 3.1 |
| Bio_09 | Mean temperature of driest quarter | 2.9 | 5.1 |
| Bio_04 | Temperature seasonality | 2.4 | 10.5 |
| Bio_18 | Precipitation of warmest quarter | 2.4 | 6.1 |
| Bio_14 | Precipitation of driest month | 1.6 | 3.6 |
| Bio_12 | Annual precipitation | 1.4 | 3.8 |
| Bio_07 | Temperature annual range | 1.2 | 3.7 |
| Bio_17 | Precipitation of driest quarter | 1.1 | 2 |
| Bio_02 | Mean diurnal range | 0.7 | 1.4 |
| Bio_08 | Mean temperature of wettest quarter | 0.7 | 1.6 |
| Bio_06 | Min temperature of coldest month | 0.5 | 3.9 |
| Bio_15 | Precipitation seasonality | 0.4 | 2.9 |
| Bio_16 | Precipitation of wettest quarter | 0.4 | 0.6 |
| Bio_19 | Precipitation of coldest quarter | 0.4 | 1.1 |
| Bio_13 | Precipitation of wettest month | 0.3 | 0.4 |

[a]Detailed interpretation see (Hijmans et al. 2005) and http://www.worldclim.org/

[b]Percent contribution means the estimated contribution of a variable to the species distribution model, and permutation importance means the loss of model fit if a particular variable is permuted randomly on landscape (Phillips and Dudík 2008)

## 3.4 Genetic Variation in North American *S. viridis* Populations

There have been two major reports aimed at characterizing the genetic diversity within populations of *S. viridis* that inhabit North America. An early study by Wang et al. used 13 allozyme markers and 168 accessions of *S. viridis* and *S. italica* sampled from populations spanning the globe to examine genetic diversity and population structure (127 accessions are from North America) (Wang et al. 1995). A more recent study (Huang et al. 2014) used genotyping by sequencing (GBS) approach to examine the genome-wide genetic diversity of *S. viridis* and *S. italica*, with a slightly denser sampling and a more directed focus on distinct regions within North America (160 North American *S. viridis* accessions out of 217 accessions). Both studies revealed important and fundamental knowledge regarding the genetic diversity within these populations. A summary of the major conclusions from these studies is presented below and the implications of how this understanding will guide future research are also discussed. However, it is worth noting that the sampling

density of both studies encompasses only a limited representation of total regional germplasm due to the large distribution range of this species. It is likely that what is currently known represents merely a general outline of a more complex pattern of genetic diversity in North American *S. viridis* populations.

### *3.4.1 Genetic Diversity*

To examine the genetic diversity captured in North American populations of *S. viridis*, Wang et al. compared summary statistics such as allelic richness and percent polymorphic loci of *S. viridis* to those other weedy species (Wang et al. 1995). They concluded that the allozyme diversity in *S. viridis* is comparable but slightly lower than other weeds. Watterson's θ (Watterson 1975) is one of the most commonly used measures of molecular genetic diversity to describe genomic sequence data and provides an estimate of expected heterozygosity in a population. It is calculated using multiple DNA sequences alignments of multiple genetic markers within a species. Due to the relatively high missing data rate associated with GBS approaches (Beissinger et al. 2013), it is difficult to calculate a precise estimate of θ directly. A reanalysis of the data from a recent GBS study (Huang et al. 2014) shows that the genome wide estimate of Watterson's θ for *S. viridis* is about $2.5 \times 10^{-3}$/base pair (bp) if one considers only the high-confidence SNPs (as listed in Huang et al. 2014), or about $1.6 \times 10^{-2}$/bp if all SNP calls are included. Neither estimate is particularly precise, due to omission of true SNPs by stringent SNP filtering or the inclusion of false positive SNPs to make such estimates. However, they provide lower and upper boundaries of the likely true estimate of $\theta$. This range is, generally speaking, comparable to other crop relatives with extensive genome-wide surveys. For example, in the maize relative teosinte, the average observed θ is approximately $1.2 \times 10^{-2}$/bp (Wright et al. 2005), and in the rice relative *Oryza rufipogon* it is $9.62 \times 10^{-3}$/bp (Huang et al. 2012a). Also, genetic analysis from both allozyme and SNP data show *S. viridis* is primarily an inbreeding species, with a selfing rate of about 96 % (Huang et al. 2014), in agreement with the traditional species descriptions.

### *3.4.2 Population Genetic Structure and Linkage Disequilibrium*

Initial assessments of genetic population structure, based on results from an allozyme diversity survey, have shown little overall genetic differentiation between Asian-European and North American *S. viridis* populations and a deep population divergence in central North America, which occurs geographically at latitude 43.5°N (Wang et al. 1995). These findings are generally confirmed and extended by a more recent study (Huang et al. 2014). In particular, the more recent study

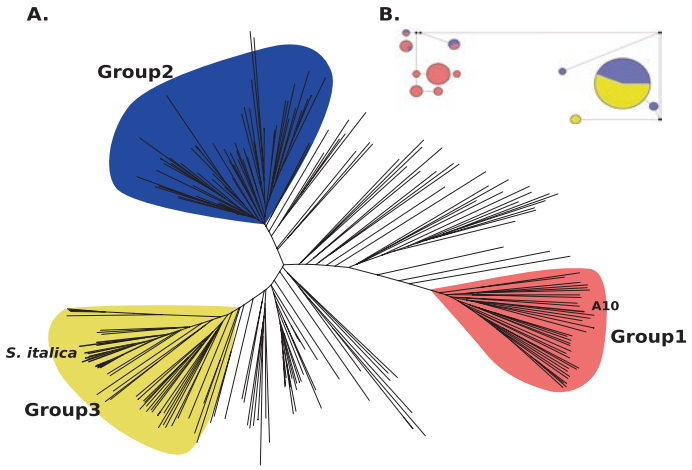


**Fig. 3.4** Genetic diversity of North American *S. viridis*. (**a**) A neighbor joining tree of a subset collection of 249 individuals of *S. viridis* and *S. italica* accessions using 13,819 randomly chosen SNP markers. Shadow area show three distinct genetic groups. The reference line A10 and *S. italica* accessions are also marked. (**b**) A median joining haplotype network of *PEPC* gene. Length of lines roughly corresponding to number of mutational changes between haplotypes connected. The colors in each pie chart correspond to the proportion of accessions that belong to certain genetic groups in **a** (highest probability of assignment)

identified two genetic clusters (Group 1 and Group 2) separated across a north–south geographic boundary and at least one additional genetic group that contained some *S. italica* introgression. Although the majority of these accessions are from China, a few North American accessions also show a genomic signature of *S. italica* introgression (Fig. 3.4a). Second, collections from the Pacific Northwest area, which was under-sampled in the earlier study, show a low level of differentiation relative to the "southern" group (Group 2), compared to their sharp distinction from the "northern" group (Group 1). Third, although the "typical" individuals from each subpopulation are very distinct from each other, there are many individuals with different degrees of admixture among these three groups. The admixed accessions tend to be positioned between their corresponding genetic groups, in both genetic diversity space and geographical space.

The pattern of linkage disequilibrium (LD) decay along physical distances in North American *S. viridis* is apparently highly affected by the population structure (Huang et al. 2014). The LD decay rate provides a general estimate of the expected resolution and the required marker density for GWAS. In *S. viridis*, LD decays rapidly, often within 45 kb. This is slightly larger but comparable to *Arabidopsis thaliana*, in which a few successful GWAS studies have been conducted (Fournier-Level et al. 2011; Hancock et al. 2011; Long et al. 2013) and where the gene density is much higher. The large number of admixed individuals among different subpopulations of *S. viridis* likely has contributed to this fast LD decay. However, among most accessions of each subpopulation, the LD decay is near or above 100 kb. This is especially true in Group 2 where there are haplotype blocks spanning long physical distances.

### 3.4.3 Biogeography and Demographic History

It has long been assumed that *S. viridis* has been introduced to North American by humans, likely during the post-Columbian era (Dekker 2003; Rominger 1962). A bottleneck followed by a range expansion scenario is a typical hypothesized demographic scenario for recently introduced invasive weeds such as *S. viridis*. Bottlenecks lead to a strong reduction of genetic diversity, and rapid range expansion expands the few haplotypes across large geographic area. Over a short window of time, mutation rate is insufficient to recover the lost genetic diversity, and instead results in a low frequency of new mutations. However, the genome-wide pattern of polymorphism in North American *S. viridis* populations provides multiple lines of evidence at variance with this hypothetical scenario (Huang et al. 2014). First, the overall genetic diversity in North American *S. viridis* populations is high. There are three deeply diverged subpopulations that coexist. Secondly, these subpopulations show genetic proximity to *S. viridis* accessions that are found in different parts of Asia and Europe, which is an indication of multiple rather than a single introduction event. Finally, a single dominating long haplotype surrounded by haplotypes that differ in a few low frequency SNPs is not observed. Given this evidence, we suggest that the North American *S. viridis* population is not reflective of a single recent introduction, but rather it is likely that multiple introductions from Asia and Europe have occurred at multiple times, with some likely occurring pre-Columbian.

## 3.5 Expanding Germplasm Collections for *S. viridis*

Given the basic properties of North American *S. viridis* natural populations outlined above, one application of this knowledge is to devise a diversity panel of germplasm for GWAS. There are a few reasons why the existing collection described in Huang et al. (2014) is not ideal for GWAS. Notably, the current sample size is smaller than successful association panels in other species like *A. thaliana* (Fournier-Level et al. 2011), rice (Huang et al. 2010, 2012b), sorghum (Morris et al. 2013), and the domesticated relative of *S. viridis*, *S. italica* (Jia et al. 2013b). In addition, strong population structure generates large LD blocks decreasing the resolution and effective sample size. The detection power may also be affected by the effects of the very different genomic backgrounds of subpopulations. For example, Si005789m.g is a gene coding a phosphoenolpyruvate carboxylase (PEPC), which is one of the key $C_4$ enzymes. Genetic variation in the PEPC gene may confer variation in photosynthetic efficiency across different lines. However, a haplotype network of this gene (Fig. 3.4b) reveals two major haplotype groups associated with this gene, and the long disjunction largely corresponds to the known population structure. Thus, it is likely that phenotypic effect difference between the two haplotype groups cannot be detected using GWAS, as effects of population structure need to be removed (Brachi et al. 2011).

Accordingly, it is clear that further expansion of existing *S. viridis* collections is imperative and will be a fruitful endeavor. Presently, we have assembled a *S. viridis* collection of more than 400 accessions covering a large geographic area across the US, a proportion of which are shown in Fig. 3.1. These new collections are more evenly distributed geographically and cover many large geographic gaps present within previous collection efforts. Expanding this collection will enable detection of more subtle population structure, as well as identifying additional admixed individuals and new genetic variants, all of which will facilitate construction of more powerful diversity panels that can be used in GWAS. Preliminary genetic diversity surveys of these new accessions, based on shotgun genome sequencing, reveal similar patterns of genetic diversity to the previous GBS study. For example, the three genetic groups largely remain in the expanded collection (Fig. 3.4a), and across the entire sample, the LD blocks are generally localized (Fig. 3.5), indicating a rapid LD decay in the total sample. In subpopulations, the LD decay is inevitably slower than the total population, and there are occasionally long distance LD blocks over 200 kb. However, large LD blocks spanning megabases of the genome originally observed in Group 2 in the GBS study (Huang et al. 2014) start to break down, probably due to the additional recombination events introduced by the new collections (Huang et al. unpublished data).

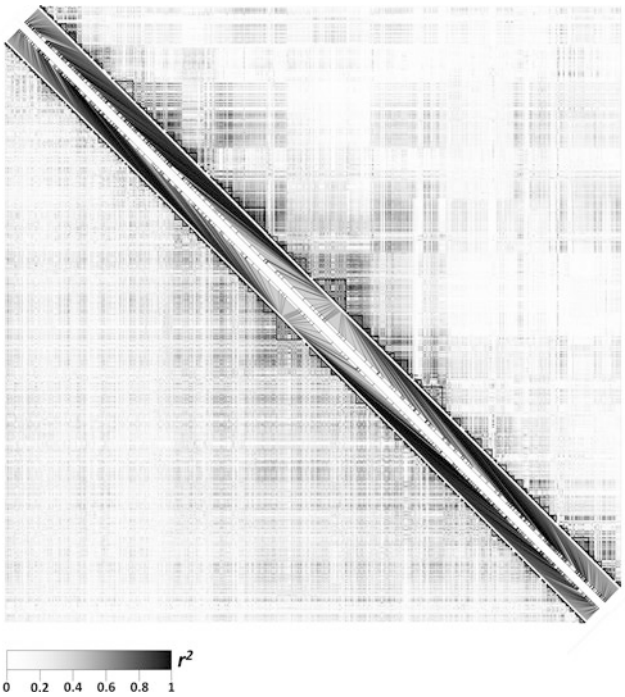


**Fig. 3.5** Pairwise linkage disequilibrium of scaffold 4. *Lower diagonal* shows the result of a subset collection of 249 accessions, and the *upper diagonal* shows the result of only individuals who belong to Group 2 (probability of assignment >90 %). *Darker colors* indicate stronger linkage disequilibrium. *Lines* in the diagonal denote relative physical locations of markers

The germplasm collection for *S. viridis* is still expanding. In the near future, accessions will be deposited in public germplasm banks for broader community access. Presently, a proportion of these collections have been submitted to USDA-GRIN by Jiang et al. for further seed propagation and distribution (www.ars-grin.gov/npgs). In addition, multiple directed crosses have been made among some of the distantly related accessions in these collections to generate new recombinant inbred lines (RILs, Chap. 18). These RIL populations can be used for further linkage mapping as well as combined for nested association mapping (NAM) designs (Buckler et al. 2009; McMullen et al. 2009). Finally, most of the germplasm collections have been or will be whole genome sequenced in deep coverage (20–30X), allowing future GWAS with various designs.

## 3.6 Opportunities for Gene Discovery

Preliminary observations indicate extensive phenotypic variation present within these new germplasm collections. One category of traits exhibiting substantial quantitative variation is panicle morphology (Fig. 3.2). Individual traits such as panicle size, spikelet branching pattern, number of spikelets, bristle color, and length are potentially determined by separate genetic mechanisms (Jia et al. 2013b; Doust et al. 2005). Revealing the underlying genes associated with these traits may have significance in crop improvement, as they may influence both seed yield and production quality in foxtail millet and other panicoid grasses. Many traits that influence plant morphology and architecture, such as plant height, tiller number, and branching, are also highly variable across this collection. Variability within these traits is inevitably influenced by environmental factors, and as such, will require increased replication and overall effort to map (Doust et al. 2004; Doust and Kellogg 2006). Finally, other important traits which are significant in bioenergy crop development, such as flowering time (Mauro-Herrera et al. 2013), photosynthetic efficiency, and biochemical compositions (Petti et al. 2013), are also likely to be highly variable among accessions. Nonetheless, GWAS and linkage mapping will be primary tools utilized for gene discovery, all of which is made possible due to these germplasm collections.

In recent years, apart from the "traditional" phenotypes, GWAS has also been employed to examine local adaptation. For example, in *A. thaliana*, associations between various climatic variables and genotypes have been examined, and signals of local adaptation have been detected using enrichment of nonsynonymous mutations in the tails of distributions (Hancock et al. 2011). There have also been more direct common garden experiments to examine climate on fitness to reveal genes responsible for local adaptation (Fournier-Level et al. 2011). Studies performed in *Populus trichocarpa* identified candidate genomic regions that show signals of local adaptation that also tend to associate with known traits with adaptive significance such as bud flush (Evans et al. 2014). Similar strategies are likely applicable to *S. viridis*, due to its wide distribution and ability to colonize diverse local environments. Genes associated with certain environmental variables (such as soil pH, precipitation,

and temperature; Fig. 3.3) may confer local adaptations to particular environments. Many of the germplasm collections have exact collecting coordinates, thus the climatic variables can be easily extracted from published interpolated climate layers (Hijmans et al. 2005). In addition, for a large proportion of these accessions, soil samples from the A horizon at each collection site have been collected and characterized. Quantitative attributes, including soil reactivity (pH), electrical conductivity, and elemental profile, are currently being analyzed (Fig. 3.3). Early comparison of values determined empirically and those downloaded from the Soil Survey Geographic Database (SSURGO) (http://websoilsurvey.nrcs.usda.gov) find significant correlation between these datasets ($r$=0.2–0.4).

Finally, as GWAS could be limited by the available samples and population structure, the ongoing deep sequencing project provides an opportunity for an alternative approach. Bulk segregant analysis (BSA) using high throughput sequencing has been applied in different study systems (Takagi et al. 2015; Yuan et al. 2013) and is a powerful method to discover genes responsible for phenotypic variations in induced mutant populations (Takagi et al. 2015). Similar methodology could potentially be used between two accessions with distinct phenotypes following protocols provided in Chap. 18 of this book. Since genotypes of all accessions are revealed through deep whole-genome sequencing, crossing schemes can be considered to minimize the genetic divergence between two parental accessions while investigating prominent phenotypic differences. The rapid life cycle and ease of propagation of *S. viridis* could make this method feasible for widespread adoption, and the sequenced parental genomes provide anchor points for quality controls. In the long run, the resulting RIL population can be combined with the diversity panel to expand the panel for a NAM population.

## 3.7 Conclusion and Future Directions

Studies characterizing the genetic diversity in North American *S. viridis* populations reveal both great potential and several major challenges for future studies. With the rapidly expanding germplasm collections and enormous sequencing efforts, knowledge of the genetic diversity in North American *S. viridis* is accelerating our ability to genetically dissect critical questions in plant biology and agriculture. In the future, one major challenge will be how to best utilize these various resources to construct a practical-sized GWAS panel that retains high mapping resolution and QTL detection power. Efficient characterization of this diversity panel will require continued development of high throughput automated systems to facilitate large scale high-precision phenotyping. Finally, collaborative efforts from various disciplines such as morphology, biogeography, physiology, and biochemistry will be combined with genetics to dissect comprehensive traits of broad interest, such as $C_4$ photosynthesis, nutrient use efficiency, and abiotic stress tolerance in *S. viridis*. It is our goal to translate the knowledge obtained from this model system to accelerate genetic improvement of economically important panicoid grass crops making them more productive and sustainable.

**Acknowledgment** Pu Huang is funded by United States National Science Foundation, Award number IOS-1127017 and United States Department of Energy, Award number DE-SC0008769. Max Feldman is funded by United States Department of Energy's Biological and Environmental Research Program Grant number ER65472. The authors thank Dr. Thomas Brutnell from Donald Danforth Plant Science Center for helpful comments on the manuscript.

## References

Barton L, Newsome SD, Chen F-H, Wang H, Guilderson TP, Bettinger RL. Agricultural origins and the isotopic identity of domestication in northern China. Proc Natl Acad Sci. 2009;106(14):5523–8.

Beissinger TM, Hirsch CN, Sekhon RS, Foerster JM, Johnson JM, Muttoni G, et al. Marker density and read depth for genotyping populations using genotyping-by-sequencing. Genetics. 2013;193(4):1073–81.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant *Setaria*. Nat Biotechnol. 2012;30(6):555–61.

Bergelson J, Roux F. Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. Nat Rev Genet. 2010;11(12):867–79.

Brachi B, Morris GP, Borevitz JO. Genome-wide association studies in plants: the missing heritability is in the field. Genome Biol. 2011;12(10):232.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X-G, et al. *Setaria viridis*: a model for $C_4$ photosynthesis. Plant Cell. 2010;22(8):2537–44.

Brutnell TP, Bennetzen JL, Vogel JP. Brachypodium distachyon and Setaria viridis: model genetic systems for the grasses. Annu Rev Plant Biol. 2015;66:465–85.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, et al. The genetic architecture of maize flowering time. Science. 2009;325(5941):714–8.

Dekker J. The foxtail (Setaria) species-group. Weed Sci. 2003;51(5):641–56.

Délye C, Wang T, Darmency H. An isoleucine-leucine substitution in chloroplastic acetyl-CoA carboxylase from green foxtail (Setaria viridis L. Beauv.) is responsible for resistance to the cyclohexanedione herbicide sethoxydim. Planta. 2002;214(3):421–7.

Doust AN, Kellogg EA. Effect of genotype and environment on branching in weedy green millet (Setaria viridis) and domesticated foxtail millet (Setaria italica) (Poaceae). Mol Ecol. 2006;15(5):1335–49.

Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA. Genetic control of branching in foxtail millet. Proc Natl Acad Sci U S A. 2004;101(24):9045–50.

Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA. The genetic basis for inflorescence variation between foxtail and green millet (Poaceae). Genetics. 2005;169(3):1659–72.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL. Foxtail millet: a sequence-driven grass model system. Plant Physiol. 2009;149(1):137–41.

Elith J, Leathwick JR. Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst. 2009;40(1):677.

Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, et al. Population genomics of Populus trichocarpa identifies signatures of selection and adaptive trait associations. Nat Genet. 2014;46(10):1089–96.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. A map of local adaptation in Arabidopsis thaliana. Science. 2011;334(6052):86–9.

Gore MA, Chia J-M, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, et al. A first-generation haplotype map of maize. Science. 2009;326(5956):1115–7.

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, et al. Adaptation to climate across the Arabidopsis thaliana genome. Science. 2011;334(6052):83–6.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005;25(15):1965–78.
Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, et al. Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet. 2010;42(11):961–7.
Huang P, Molina J, Flowers JM, Rubinstein S, Jackson SA, Purugganan MD, et al. Phylogeography of Asian wild rice, Oryza rufipogon: a genome-wide view. Mol Ecol. 2012a;21(18):4593–604.
Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet. 2012b;44(1):32–9.
Huang P, Feldman M, Schroder S, Bahri BA, Diao X, Zhi H, et al. Population genetics of Setaria viridis, a new model system. Mol Ecol. 2014;23(20):4912–25.
Jia G, Shi S, Wang C, Niu Z, Chai Y, Zhi H, et al. Molecular diversity and population structure of Chinese green foxtail [Setaria viridis (L.) Beauv.] revealed by microsatellite analysis. J Exp Bot. 2013a;64(12):3645–56.
Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (Setaria italica). Nat Genet. 2013b;45(8):957–61.
Jiang H, Barbier H, Brutnell T. Methods for performing crosses in Setaria viridis, a new model system for the grasses. J Vis Exp. 2013;80:e50527.
Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, et al. Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. Nat Genet. 2013;45(8):884–90.
Mauro-Herrera M, Wang X, Barbier H, Brutnell TP, Devos KM, Doust AN. Genetic control and comparative genomic analysis of flowering time in Setaria (Poaceae). G3 Genes Genom Genet. 2013;3(2):283–95.
McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, et al. Genetic properties of the maize nested association mapping population. Science. 2009;325(5941):737–40.
Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, et al. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc Natl Acad Sci. 2013;110(2):453–8.
Petti C, Shearer A, Tateno M, Ruwaya M, Nokes S, Brutnell T, et al. Comparative feedstock analysis in Setaria viridis L. as a model for $C_4$ bioenergy grasses and Panicoid crop species. Front Plant Sci. 2013;4:181.
Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 2008;31(2):161–75.
Rominger JM. Taxonomy of Setaria (Gramineae) in North America. Illinois Biol Monogr. 1962;29:1–132.
Sebastian J, Wong MK, Tang E, Dinneny JR. Methods to promote germination of dormant Setaria viridis seeds. PLoS One. 2014;9(4):e95109.
Takagi H, Tamiru M, Abe A, Yoshida K, Uemura A, Yaegashi H, et al. MutMap accelerates breeding of a salt-tolerant rice cultivar. Nat Biotechnol. 2015;33(5):445–9.
Wang R, Wendel JF, Dekker JH. Weedy adaptation in Setaria spp. I. Isozyme analysis of genetic diversity and population genetic structure in Setaria viridis. Am J Bot. 1995;82:308–17.
Wang C, Chen J, Zhi H, Yang L, Li W, Wang Y, et al. Population genetics of foxtail millet and its wild ancestor. BMC Genet. 2010;11(1):90.
Watterson G. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 1975;7(2):256–76.
Wilson P, Streich J, Borevitz J. In: Genomic diversity and climate adaptation in Brachypodium. New York: Springer; 2015. p. 1–21. doi:10.1007/7397_2015_18.
Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, et al. The effects of artificial selection on the maize genome. Science. 2005;308(5726):1310–4.
Yuan Y-W, Sagawa JM, Di Stilio VS, Bradshaw HD. Bulk segregant analysis of an induced floral mutant identifies a MIXTA-like R2R3 MYB controlling nectar guide formation in Mimulus lewisii. Genetics. 2013;194(2):523–8.

# Chapter 4
# Origin and Domestication of Foxtail Millet

**Xianmin Diao and Guanqing Jia**

**Abstract** Among the more than 100 species in *Setaria*, *S. macrostachya*, *S. pumila,* and foxtail millet (*S. italica*) cereals were domesticated by human beings. However, only foxtail millet became a worldwide crop, contributing greatly to the development of Chinese civilization and remaining as a staple cereal in arid and semi-arid regions. Green foxtail is the ancestor of cultivated foxtail millet and both can be regarded as the same species. Archeological evidence indicates that the domestication of foxtail millet from green foxtail probably began around 16,000 YBP, was a recognized crop around 9000–10,000 YBP, and became popular in Northern China at about 5000–6000 YBP, then spread to other parts of the world. Although there has been some controversies over whether the domestication of foxtail millet has occurred more than once, recent molecular data and archeological evidence suggest a single domestication event.

**Keywords** Foxtail millet • Green foxtail • Domestication • Setaria

## 4.1 Origin and Spread of *Setaria*

The genus *Setaria* comprises several subgenera with reportedly 125 species distributed worldwide (Hubbard 1915; Rominger 1962), although recent data suggests that there are only 99 species (Kellogg 2015; Kellogg et al. 2009). Although *Setaria* was first described more than 260 years ago by Linnaeus (1753), the origin of this group of species remains unclear. It has been proposed that the ancestor of the current Eurasian *Setaria* spp. originated from Africa, and that the first species invading Eurasia from Africa was probably green foxtail (*Setaria viridis*) or a *S. viridis*-like diploid annual, which then gave rise to other species of the Setaria grass (Rominger 1962; Stapf and Hubbard 1930). The reason for this conjecture might be that most *Setaria* species are located in Africa,

X. Diao (✉) • G. Jia
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences,
No. 12, Zhongguancun South St., Haidian District, Beijing 100081, People's Republic of China
e-mail: diaoxianmin@caas.cn



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_4

and it is the only theory on the origin of *Setaria* to date. However, recent molecular phylogenies emphasize that Setaria is polyphyletic, with strongly supported African and Asian, as well as South American clades, but with no clear pattern of relationship between them or with other recognized taxa, making distributional hypotheses problematic (Kellogg et al. 2009) (Chap. 1).

It is widely accepted that green foxtail and foxtail millet (*S. italica*) are native to Eurasia, and that green foxtail now has a worldwide distribution, yet we know very little about how it expanded into its current distribution. One hypothesis (Dekker 2003) suggests that *Setaria* expanded from Africa to its current situation in the northern hemisphere in five major phases, including the spread of green foxtail into North America as a post-Columbian invasion event. However, recent single nucleotide polymorphism (SNP) analyses of the diversity and population structure of green foxtail samples from North America suggest that there have been multiple introductions into North America from distinct gene pools in China, Central Asia, and Europe (Huang et al. 2014, Chap. 3). The origin and spread of green foxtail still needs more study.

## 4.2 *Setaria* Species Used as Cereals or Fodder

Several *Setaria* species have been domesticated and cultivated as cereals for human consumption. Foxtail millet was first domesticated in China and then became a cereal cultivated throughout Eurasia. It remains a major crop in arid and semi-arid regions of China and India to this day. Yellow foxtail, *S. pumila* (Poir.) Roem. & Schult. (syn. *Setaria glauca*, *Setaria lutescens* F. T. Hubbard) was cultivated and domesticated as a grain crop in Eastern and Western Ghats, India, and its grains were harvested from wild plants elsewhere (Watt 1908; Datta and Banerjee 1978; Rao et al. 1987; de Wet 1992; Klmata et al. 2000; Dekker 2003; Austin 2006). *S. macrostachya* was used and domesticated as a cereal crop in Mexico before the domestication of maize (Callen 1967; Smith 1967; Felger and Moser 1985; Austin 2006). In addition to foxtail millet, seeds of yellow foxtail, *S. macrostachya*, and other *Setaria* species were gathered from wild plants as grain food, such as *Setaria liebmannii* E. Fournier in Mexico, *Setaria pallide-fusca* (Schumacher) Stapf and C. E. Hubbard in Africa, *Setaria palmifolia* (J. König) Stapf in Asia, and *Setaria sphacelata* (Schumacher) M. B. Moss ex Stapf and C. E. Hubbard in Africa. Other *Setaria* species were used as fodder, such as *Setaria intermedia* Rothex Roemer and Schultes (syn. *Setaria tomentosa* (Roxburgh) Kunth) in India and *S. pallide-fusca* in Nepal (Austin 2006).

Although foxtail millet, yellow foxtail, and *S. macrostachya* were domesticated as cereals, only foxtail millet became a worldwide crop, contributing greatly to the development of human civilization in Eurasia for more than 10,000 years. It remains a staple grain in arid and semi-arid region of China and India. In this chapter, we focus on the origin and domestication of foxtail millet.

## 4.3 Domestication of Foxtail Millet

### *4.3.1 Genetic Evidence Supporting Green Foxtail as the Ancestor of Foxtail Millet*

In the eighteenth-century, Linnaeus (1753) classified green foxtail and foxtail millet into the genus *Panicum*, but named them as two different species, *Panicum viride* and *Panicum italica*, respectively. Later they were transferred to the genus *Setaria*, and their botanical names were changed to *S. viridis* and *S. italica*, respectively, remaining as two independent species (Austin 2006). In the twentieth-century, many researchers observed morphological and cytological similarities between green foxtail and foxtail millet, found that they produce fertile hybrids, and noticed that they share continuous and overlapping genetic variation (Darmency et al. 1987a; Li et al. 1935, 1942; Kihara and Kishimoto 1942; Rao et al. 1987; Takahashi and Hoshino 1934; Till-Bottraud et al. 1992; Willweber-Kishimoto 1962). An isozyme analysis demonstrated that green foxtail and foxtail millet share many similarities at the biochemical level (Wang et al. 1995a, b). Using random amplified polymorphic DNA (RAPD) markers, Li et al. (1998) found that green foxtail and foxtail millet collected from different regions formed mixed clusters. In later studies, restriction fragment length polymorphism (RFLP), inter-simple sequence repeat (ISSR), and transposon display analyses confirmed that green foxtail is the ancestor of foxtail millet (Wang et al. 1998; Li et al. 2012; Hirano et al. 2011). After the release of the reference genome of foxtail millet (Bennetzen et al. 2012; Zhang et al. 2012), the first *de novo* assembly of the genome sequence of green foxtail was produced. Sequence comparisons between green foxtail and foxtail identified only small differences between the two and confirmed that green foxtail is the ancestor of foxtail millet (Jia et al. 2013).

These findings led to the theories that green foxtail and foxtail millet are actually the same species, and that the difference in stature between them reflects the differentiation of wild and domesticated types. Rao et al. (1987) suggested that both taxa should be considered as subspecies of *S. italica* (green foxtail, *S. italica*, subsp. *viridis*; foxtail millet, *S. italica*, subsp. *italica*). However, this nomenclature implied the primacy of the crop over its wild *Setaria* ancestors, and Dekker (2003) suggested that the two sub-species be named foxtail millet (*S. viridis* subsp. *italica*) and green foxtail (*S. viridis* subsp. *viridis*). No one naming convention has prevailed, and all combinations are found in the recent literature.

### *4.3.2 Wild-Weed-Crop Complexes in Setaria*

The domestication of wild plant species to produce cultivated crops is thought to begin with gathering seeds from the wild, before progressing to deliberate planting and gathering. Gradual genetic changes in the wild plant lead to domestication-related

characteristics that are desirable for human cultivation, such as suppression of seed shattering and selection of appropriate flowering time (named the "domestication syndrome"). At this stage, the plant is considered domesticated (Dekker 2003; Austin 2006). Foxtail millet, as well as green foxtail, is a self-pollinating species with 0.3–4 % outcrossing rates under natural conditions (Li et al. 1945; Till-Bottraud et al. 1992). The outcrossing rate can be as high as 7 % depending on the geography and climatic conditions during growth and flowering (summarized by Liang and Quan (1997)). A study on the dispersal of foxtail millet pollen demonstrated that viable pollen can be blown over a long distance depending on the weather conditions (Wang et al. 2001). The self-pollination and outcrossing characters of green foxtail and foxtail millet have also been verified by DNA marker analysis (Wang et al. 2010). The selfing rate of green foxtail was estimated at 96 % based on SNP data and 90 % based on simple sequence repeat (SSR) data (Huang et al. 2014; Jia et al. 2013), compared with 98 % for foxtail millet, as estimated from SSR data (Wang et al. 2012).

Spontaneous hybridizations between the wild parent (green foxtail) and the crop (foxtail millet) have occurred in the fields ever since foxtail millet was domesticated. Furthermore, their hybrids are fertile. Consequently, traits of weedy green foxtail and domesticated foxtail millet have been exchanged in millet fields since first cultivation (Darmency et al. 1987b; Harlan 1965; Harlan et al. 1973; Dekker 2003). Genomic alleles for wild characters from green foxtail have been largely eliminated from cultivated foxtail millet by intensive human selection. However, genes introgressed into the wild green foxtail might be retained if they have advantages for natural selection, as evidenced by the many weedy types with morphological similarities to foxtail millet that arise around foxtail millet fields. Human selection of foxtail millet has focused on tall plant stature and robust growth for higher grain yield. The tall plant stature character has transferred from foxtail millet into green foxtail, leading to the emergence of giant green foxtail (*S. viridis* var. major (Gaudin) Pospichal) around farmed fields (Pohl 1951, 1966; Rominger 1962; Darmency et al. 1987b). This is an example of a wild-weed-crop complex (Darmency et al. 1987b; Rao et al. 1987), where the partial reproductive barriers between crop and weed variants become an advantage for the weed when a new variant acquires weedy traits and invades the human-managed habitat (Darmency et al. 1987b). Genetic evidence also suggests that more recent or ongoing cross-pollination is occurring in foxtail millet production regions (Wang et al. 2010).

### *4.3.3 Archaeological Evidence for the Domestication and Spread of Foxtail Millet*

In ancient China, agriculture was based on the domestication of broomcorn or common millet (*Panicum milieaceum*) and foxtail millet (*S. italica*) in Northern China and rice (*Oryza sativa*) in Southern China. Qinling Mountain marks the boundary between Northern and Southern China. Because the political center of ancient

China was generally located in the north, common millet and foxtail millet greatly contributed to the development of Chinese civilization, as evidenced in many studies (He and Hui 2015; Shelach 2000; Sigaut 1994; Wei 1994). Archaeological discoveries have provided incontrovertible and direct evidence for the origins and domestication of crop plants, including foxtail millet.

The first phase of foxtail millet domestication occurred from approximately 23,000 YBP to around 9000 YBP. No intact foxtail millet grains corresponding to this phase have been found, but there is ample evidence of plant starches and stone tools for processing green foxtail and/or foxtail millet and specific evidence of foxtail millet starch (Liu et al. 2013; Yang et al. 2012). The oldest stone tools for processing green foxtail were found at the Shizitan site in southern Shanxi, dated back to 23,000 to 19,500 YBP (Liu et al. 2013). Further archeological evidence was found at the Xiachuan site in southern Shanxi Province, dating back to 16,000 YBP (Wei 1994). Archeological remains have also been found at the Nanzhuangtou site (older than 11,000 YBP) and the Donghulin site (10,150 to 9500 YBP) near Beijing, and the Shizitan site (9600–9000 YBP) near the Yellow River in southern Shanxi Province (Liu et al. 2011). Because foxtail millet grains from this phase have never been found, it was deduced that green foxtail was collected from the wild as food and the domestication of foxtail millet began around this time (Yang et al. 2012).

The second phase of foxtail millet domestication took place between 9000 and 6000 YBP, in the middle phase of the Chinese Neolithic Age. The oldest foxtail millet grains found to date were retrieved from the Donghulin site in Beijing, and dated back to 11,000 to 9000 YBP (Zhao 2014). Foxtail millet grains were also found at the Zhangmatun site in Shandong Province, dating back to 9000 to 8500 YBP (Wu et al. 2014). A phytolith analysis positively identified foxtail millet grains found at the Cishan site in Wu'an County, southern Hebei Province, which dated back to 8700 YBP (Lü et al. 2009). However, there has been some discussion about the dating of those grain samples. Many older intact carbonized foxtail millet grains have been found in Northern China, mainly in the middle and upper Yellow River region, dating to between 8500 and 6000 YBP (Fig. 4.1). Sixty carbonized foxtail millet grains were among nearly 1500 grains of common millet and foxtail millet found at the Xinglonggou site in east Inner Mongolia, dating back to 8000 to 7500 YBP. Other evidence of foxtail millet from this phase includes remains found at sites of the Peiligang culture in Henan Province (8500 to 7000 YBP) and the Dadiwan-Laoguantai culture in Gansu Province (7800 to 7300 YBP). From those finds, we know that green foxtail had already been domesticated to produce the larger-grained foxtail millet, and that foxtail millet had already spread throughout a wider area in Northern China, mainly around the Yellow River region. One character of agriculture in this phase in Northern China is the co-cultivation of common millet and foxtail millet. At most archeological sites in Northern China, common millet grains far outnumber foxtail millet grains, confirming that common millet was the main type cultivated during that period.

The third phase of foxtail millet development was its expansion after domestication, from approximately 6000 YBP. Carbonized foxtail millet remains have been recovered from hundreds of archaeological sites in China, with large quantities found at many sites (Fig. 4.1). Typical sites of this phase are the mid- and late-Yangshao

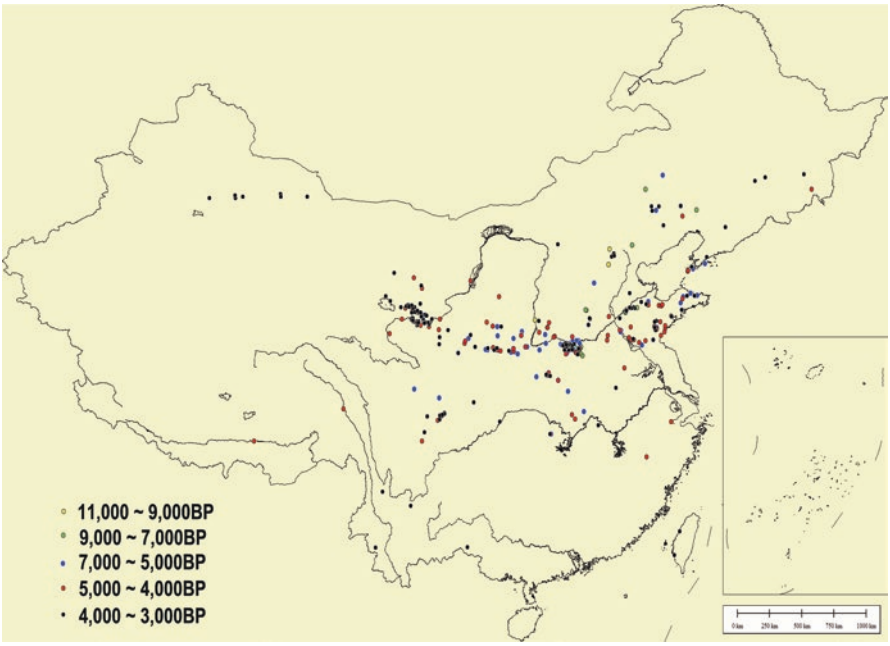


**Fig. 4.1** Distribution of archeological sites in China where foxtail millet remains have been identified. Prepared by He Keyang and Lu Houyuan. Different colors indicate time periods (age) of various archeological sites. *Yellow* and *green* points represent the oldest sites and the earliest foxtail millet remains. These sites are mostly located in the Yellow River region in Northern China, indicating that this was where foxtail millet was first domesticated

Culture sites (6000–5000 YBP) located in the middle of the Yellow River region in Henan Province. Foxtail millet remains were found at the Chengtoushan site (6000 YBP) in the middle region of the Yangtze River in Hunan Province, south of the Yellow River region where foxtail millet was first domesticated. Then, it spread from Hunan to the Chengdu Plain in Sichuan Province along the Yangtze River. This route of expansion was confirmed by the carbonized foxtail millet found at the Haxiu and Yingpanshan sites (5300–4500 YBP). From Sichuan, foxtail millet spread to the Tibet–Qinghai Plateau, as verified by remains found at the Kanuo site (5500–4200 YBP) (Lü et al. 2005, 2014). This scenario explains foxtail millet expansion to the south west of China. The prevalence of foxtail millet after its domestication in the middle of the Yellow River region has been verified by the presence of many millet grains at archeological sites of the Longshan culture (4500 YBP) in Shandong Province, east of the Yellow River region; and in those of the Hongshan culture (5500–500 YBP), north of the Yellow river in Inner Mongolia.

After its domestication, foxtail millet spread surprisingly quickly. Foxtail millet grains have been recovered from archeological sites in Eastern Siberia including Krounovka-1 (5550–5350 YBP), Zaisanovka-7 (5500–5000 YBP), Zaisanovka-1 (4600–4400 YBP), and Novoselische-4 (4500–4050 YBP) (Sergusheva and Vostretsov 2009). It spread to Korea around 5500 YBP, as verified by its presence at

the Chulmun site (5500 YBP) in southern Korea (Lee 2000). The oldest foxtail millet grains in Japan were identified at the Usujiri site in Hokkaido, dating back to 4000 YBP (Crawford 1992). Foxtail millet remains have been found at several archeological sites in Xinjiang Province, most dating back to 4000 to 3000 YBP (Fig. 4.1). The archeological evidence suggests that foxtail millet spread from Xinjiang to Europe (Wei 1994).

Archaeological evidence from sites across Europe and Asia suggests that although the progenitor of foxtail millet, green foxtail, was widespread throughout Eurasia, carbonized foxtail millet grains dating to before 7000 YBP were found only in China (Jones 2004; Daniel and Maria 2000; Hunt et al. 2008). This provides further evidence that China was the site of foxtail millet domestication, albeit there may have been more than one domestication event for this species. It is enigmatic that there is no concrete evidence for foxtail millet until the Iron Age in the Near East (Nesbitt and Summers 1988), even though this region is geographically close to China.

The archeological evidence indicates that foxtail millet was first domesticated in the middle region of the Yellow River including Shanxi, Shannxi, Hebei, and Henan provinces. The process of domestication probably began around 16,000 YBP. Based on the size of starch grains, it changed from a grass to a crop plant around 9000–10,000 YBP. Foxtail millet became popular in Northern China at about 5000–6000 YBP. The introduction of wheat and barley into China 4000 YBP changed the agricultural position of foxtail millet, but it remained as a staple cereal for more than 5000 years in Northern China until maize (*Zea mays* L.) and sweet potato (*Ipomoea batatas* L. (Lam)) were introduced approximately 400 years ago. To this day, foxtail millet remains a staple food in some semi-arid and arid areas of Northern China. The long history of foxtail millet cultivation has greatly contributed to the development of Chinese culture and to human civilization in general. For example, the first noodles were made from millet, and noodles remain a very popular food worldwide (Lü et al. 2005, 2014). However, the archeological evidence cannot prove whether there were multiple independent origins of foxtail millet in Europe (Harlan 1975; Jusuf and Pernes 1985), central Asia (Li et al. 1995; Sakamoto 1987), and tropical Asia (Fukunaga et al. 2006). Combining the results of diverse types of analyses will provide more details about the history of foxtail millet domestication.

### *4.3.4 Monophyletic or Polyphyletic Origin of Foxtail Millet*

Cytological and hybridization studies have shown that green foxtail is the wild ancestor of foxtail millet (Li et al. 1942; Kihara and Kishimoto 1942). However, the geographical origin of domesticated foxtail millet cannot be determined from the distribution of green foxtail because the latter is found commonly in various areas of Europe and Asia. As for the geographical origin of foxtail millet, Vavilov (1926) first suggested that China was the principal center of foxtail millet domestication. Several other hypotheses have been entertained, for instance, based on archaeological evidence, Harlan (1975) suggested an independent domestication center in Europe,

because he assumed that foxtail millet could not be transferred to Europe from China in ancient times (>3000 YBP). Using an isozyme marker analysis, Jusuf and Pernes (1985) found that the genetic distance between cultivated foxtail millet and wild types of the same origin was sometimes smaller than the genetic distance between accessions of foxtail millet and green foxtail from different origins, consistent with the multiple origins hypothesis. These results were supported by hybrid weakness and partial sterility in offspring derived from intraspecific crosses between Chinese and European landraces (Croullebois et al. 1989) and a cluster analysis based on the morphological characteristics of foxtail millet (Li et al. 1995). However, it is likely that the long history of cultivation and natural selection of foxtail millet varieties adapted to local environments, and the numerous hybridizations and genetic exchanges between foxtail millet and local green foxtail, have resulted in them sharing similar genetic backgrounds and isozyme patterns. Consequently, isozyme analyses should be carefully interpreted in domestication studies, such as foxtail millet.

A second hypothesis (Li et al. 1995) suggested that the landraces from Afghanistan and Lebanon had been domesticated independently and relatively recently, because they had primitive morphological characters. This idea was supported by ribosomal DNA data (Fukunaga et al. 2006). Furthermore, Sakamoto (1987) reported that foxtail millet landraces in central Asia, Afghanistan, Pakistan, and Northwest India not only had primitive morphological traits, but also showed relatively high cross-compatibility with foxtail millet from other regions (Kawase and Sakamoto 1987). Based on those results, they suggested that the original center of foxtail millet domestication was located somewhere in those regions, and China may be the secondary center of diversification. This is the only hypothesis that places China as the secondary center of foxtail millet domestication.

Many studies based on chemical markers, such as esterase isozymes, could not reach clear conclusions about the monophyletic or multiple origins of foxtail millet (Kawase and Sakamoto 1984; Wang et al. 1995a). An amplified fragment length polymorphism (AFLP) analysis designed to detect the origin of foxtail millet failed to find any geographic structure (d'Ennequin et al. 2000). Later studies based on RFLP data (Fukunaga et al. 2002) and variations in ribosomal DNA intergenic spacer subrepeats (Fukunaga et al. 2006) helped to classify foxtail millet landraces, but could not provide further information on the foxtail millet domestication hypothesis.

Some recent studies have provided valuable clues about the domestication of foxtail millet. Using ISSR markers, Li et al. identified that both Chinese and Europe foxtail millet landraces were closely related to a few green foxtail varieties collected from the middle and upper regions of the Yellow River in Northern China, consistent with archaeological evidence from sites in those regions (Li et al. 2012). An analysis based on genome-wide TE insertion polymorphisms as DNA markers identified a clear geographical structure of foxtail millet from the worldwide collection. In that structure, almost all green foxtail accessions formed an outgroup to domesticated foxtail millet. These results strongly indicate that foxtail millet has a monophyletic origin. However, because of the small sample size of green foxtail,

the detailed geographical location of its origin remained unclear (Hirano et al. 2011). The recent resequencing of 916 foxtail millet accessions from the worldwide collection and identification of more than 0.8 million SNPs has resulted in a new phylogenetic tree. This tree shows a clear geographical structure with accessions from the same local geographic regions grouped together into the same clades, suggesting a single domestication of foxtail millet (Jia et al. 2013). Furthermore, phylogenetic analyses based on sequences around the seed-shattering gene *Sh1* showed evidence of a domestication sweep. In those analyses, all foxtail millet samples from China and other countries were clustered as star-like branches, but all green foxtail samples were grouped on long branches, indicating great diversity among wild types and a single origin of foxtail millet (Jia et al. 2013). Together, the results of these recent studies suggest that the domestication of foxtail millet is monophyletic. Analyses of more green foxtail samples from around the world would strengthen the results of these studies. Also, for more in-depth studies on the origin of foxtail millet and genetic tools to reduce the noise resulting from long-time gene exchange between green foxtail and foxtail millet are needed, such as analyses of the sequences of domestication-related genes.

## References

Austin DF. Fox-tail millets (Setaria: Poaceae)—abandoned food in two hemispheres. Econ Bot. 2006;60(2):143–58.

Bennetzen JL, Schmutz J, Wang H, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30(6):555–61.

Callen EO. The first new world cereal. Am Antiq. 1967;32(4):535–8.

Crawford G. The transitions to agriculture in Japan. In: Gebauer A, Price TD, editors. Transitions to agriculture in prehistory. Madison: Prehistory Press; 1992. p. 17–132.

Croullebois ML, Barreneche MT, Cherisey H, et al. Intraspecific differentiation of Setaria italica (L.) PB: study of abnormalities (weakness, segregation distortion, and partial sterility) observed in F1 and F2 generations. Genome. 1989;32(2):203–7.

Daniel Z, Maria H. Domestication of plants in the old world. Oxfordshire: Oxford University Press; 2000.

Darmency H, Zangre GR, Pernes J. The wild-weed-crop complex in Setaria: a hybridization study. Genetica. 1987a;75(2):103–7.

Darmency H, Ouin C, Pernes J. Breeding foxtail millet (*Setaria italica*) for quantitative traits after interspecific hybridization and polyploidization. Genome. 1987b;29(3):453–6.

Datta SC, Banerjee AK. Useful weeds of West Bengal rice fields. Econ Bot. 1978;32(3):297–310.

Dekker J. The foxtail (Setaria) species-group. Weed Sci. 2003;51(5):641–56.

d'Ennequin MLT, Panaud O, Toupance B, et al. Assessment of genetic relationships between *Setaria italica* and its wild relative *S. viridis* using AFLP markers. Theor Appl Genet. 2000;100(7):1061–6.

De Wet JMJ. The three phases of cereal domestication. In: Capman GP, editor. Grass evolution and domestication. Cambridge: Cambridge University Press; 1992. p. 176–98.

Felger RS, Moser MB. People of the desert and sea: ethnobotany of the Seri Indians. Tucson: University Arizona Press; 1985.

Fukunaga K, Wang Z, Kato K, et al. Geographical variation of nuclear genome RFLPs and genetic differentiation in foxtail millet, *Setaria italica* (L.) P. Beauv. Genet Resour Crop Evol. 2002;49(1):95–101.

Fukunaga K, Ichitani K, Kawase M. Phylogenetic analysis of the rDNA intergenic spacer subrepeats and its implication for the domestication history of foxtail millet, *Setaria italica*. Theor Appl Genet. 2006;113(2):261–9.

Harlan JR. The possible role of weed races in the evolution of cultivated plants. Euphytica. 1965;14(2):173–6.

Harlan JR. Crops and man. Madison: American Society of Agronomy and Crop Science Society of America; 1975.

Harlan JR, De Wet JMJ, Price EG. Comparative evolution of cereals. Evolution. 1973;27:311–25.

He H, Hui F. The history of millet cultivation in ancient China. Beijing: China Agricultural Science and Technology Press; 2015.

Hirano R, Naito K, Fukunaga K, et al. Genetic structure of landraces in foxtail millet (*Setaria italica* (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet. Genome. 2011;54(6):498–506.

Huang P, Feldman M, Schroder S, et al. Population genetics of *Setaria viridis*, a new model system. Mol Ecol. 2014;23(20):4912–25.

Hubbard FT. A taxonomic study of *Setaria italica* and its immediate allies. Am J Bot. 1915;2(4):169–98.

Hunt HV, Vander Linden M, Liu X, et al. Millets across Eurasia: chronology and context of early records of the genera Panicum and Setaria from archaeological sites in the Old World. Veg Hist Archaeobot. 2008;17(1):5–18.

Jia G, Huang X, Zhi H, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013;45(8):957–61.

Jones M. Between fertile crescents: minor grain crops and agricultural origins. In: Jones M, editor. Traces of ancestry: studies in honour of Colin Renfrew. Cambridge: McDonald Institute for Archaeological Research; 2004. p. 127–36.

Jusuf M, Pernes J. Genetic variability of foxtail millet (*Setaria italica* P. Beauv.). Theor Appl Genet. 1985;71(3):385–91.

Kawase M, Sakamoto S. Variation, geographical distribution and genetical analysis of esterase isozymes in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1984;67(6):529–33.

Kawase M, Sakamoto S. Geographical distribution of landrace groups classified by hybrid pollen sterility in foxtail millet, *Setaria italica* (L.) P. Beauv. Jpn J Breed. 1987;37(1):1–9.

Kellogg EA. Poaceae. In: Kubitzki K, editor. The families and genera of vascular plants. Berlin: Springer; 2015. p. 1–415.

Kellogg EA, Aliscioni SS, Morrone O, et al. A phylogeny of Setaria (Poaceae, Panicoideae, Paniceae) and related genera, based on the chloroplast gene ndhF. Int J Plant Sci. 2009;170:117–31.

Kihara H, Kishimoto E. Bastarde zwischen *Setaria italica* and *S. viridis*. Bot Mag. 1942;20:63–7.

Klmata M, Ashok EG, Seetharam A. Domestication, cultivation and utilization of two small millets, Brachiaria ramosa and Setaria glauca (Poaceae), in south India. Econ Bot. 2000;54(2):217–27.

Lee G-A. The transition from foraging to farming in prehistoric Korea. Curr Anthropol. 2000;52(S4):S307–29.

Li HW, Meng CJ, Liu TN. Problems in the breeding of millet (*Setaria italica* (L.) Beuav.). J Am Soc Agron. 1935;27(12):963–70.

Li CH, Pao WK, Li HW. Interspecific Crosses in Setaria II. Cytological studies of interspecific hybrids involving: 1, *S. faberii* and *S. italica*, and 2, a three way cross, F2 of *S. italica*×*S. viridis* and *S. faberii*. J Hered. 1942;33(10):351–5.

Li HW, Li CH, Pao WK. Cytological and genetical studies of the interspecific cross of the cultivated foxtail millet, *Setaria italica* (L.) Beauv., and the green foxtail millet, *S. viridis* L. J Am Soc Agron. 1945;37(1):32–54.

Li Y, Wu S, Cao Y. Cluster analysis of an international collection of foxtail millet (*Setaria italica* (L.) P. Beauv.). Euphytica. 1995;83(1):79–85.

Li Y, Jia J, Wang Y, et al. Intraspecific and interspecific variation in Setaria revealed by RAPD analysis. Genet Resour Crop Evol. 1998;45(3):279–85.
Li W, Hui Z, Wang Y, et al. Assessment of genetic relationship of foxtail millet with its wild ancestor and close relatives by ISSR markers. J Integr Agric. 2012;11(4):556–66.
Liang H, Quan J. Biology and botany of foxtail millet. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 100–29.
Linnaeus C. Species plantarum (vol. 1). Stockholm: Laurentii Salvii; 1753.
Liu L, Ge W, Bestel S, et al. Plant exploitation of the last foragers at Shizitan in the Middle Yellow River Valley China: evidence from grinding stones. J Archaeol Sci. 2011;38(12):3524–32.
Liu L, Bestel S, Shi J, et al. Paleolithic human exploitation of plant foods during the last glacial maximum in North China. Proc Natl Acad Sci. 2013;110(14):5380–5.
Lü H, Yang X, Ye M, et al. Culinary archaeology: millet noodles in late Neolithic China. Nature. 2005;437(7061):967–8.
Lü H, Zhang J, Liu K, et al. Earliest domestication of common millet (*Panicum miliaceum*) in East Asia extended to 10,000 years ago. Proc Natl Acad Sci. 2009;106(18):7367–72.
Lü H, Li Y, Zhang J, et al. Component and simulation of the 4,000-year-old noodles excavated from the archaeological site of Lajia in Qinghai, China. Chin Sci Bull. 2014;59(35):5136–52.
Nesbitt M, Summers GD. Some recent discoveries of millet (*Panicum miliaceum* L. and *Setaria italica* (L.) P. Beauv.) at excavations in Turkey and Iran. Anatol Stud. 1988;38:85–97.
Pohl RW. The genus Setaria in Iowa. Iowa State J Sci. 1951;25:501–8.
Pohl RW. The grasses of Iowa. Iowa State J Sci. 1966;40:341–73.
Rao KEP, De Wet JMJ, Brink DE, et al. Infraspecific variation and systematics of cultivated *Setaria italica*, foxtail millet (Poaceae). Econ Bot. 1987;41(1):108–16.
Rominger JMD. Taxonomy of Setaria (Gramineae) in North America. Illinois Biol Monogr. 1962;29:1–118.
Sakamoto S. Origin and dispersal of common millet and foxtail millet. Jpn Agr Res Q. 1987;21(22):84–9.
Sergusheva EA, Vostretsov YE. The advance of agriculture in the coastal zone of East Asia. In: Fairbaim A, Weiss E, editors. From Forages to Farmers: papers in honor of Gordon C. Hillman. Oxford: Oxbow Books; 2009. p. 205–19.
Shelach G. The earliest Neolithic cultures of Northeast China: recent discoveries and new perspectives on the beginning of agriculture. J World Prehist. 2000;14(4):363–413.
Sigaut F. Les millets en Eurasie, d'une fete populaire a des questions pour les chercheurs. Industries des Cereales. 1994;10:23–32.
Smith CE. Plant remains. In: Byers DS, editor. The prehistory of the Tehuacán Valley. Austin: University of Texas Press; 1967. p. 220–55.
Stapf O, Hubbard CK. Setaria. In: Prain D, editor. Flora of Tropical Africa, vol. 9. London: W. Clowes & Sons; 1930. p. 768–866.
Takahashi N, Hoshino T. Natural crossing in *Setaria italica* (Beauv.). Proc Crop Sci Soc Jpn. 1934;6:3–19.
Till-Bottraud I, Reboud X, Brabant P, et al. Outcrossing and hybridization in wild and cultivated foxtail millets: consequences for the release of transgenic crops. Theor Appl Genet. 1992; 83(8):940–6.
Vavilov NI. The origin of the cultivation of "primary" crops, in particular cultivated hemp. In: Vavilov N, editor. Studies on the origin of cultivated plants. Leningrad: Institute of Applied Botany and Plant Breeding; 1926. p. 221–33.
Wang R, Wendel JF, Dekker JH. Weedy adaptation in Setaria spp. I. Isozyme analysis of genetic diversity and population genetic structure in *Setaria viridis*. Am J Bot. 1995a;82(3):308–17.
Wang RL, Wendel JF, Dekker JH. Weedy adaptation in Setaria spp. II. Genetic diversity and population genetic structure in *S. glauca*, *S. geniculata*, and *S. faberii* (Poaceae). Am J Bot. 1995b;82(8):1031–9.
Wang ZM, Devos KM, Liu CJ, et al. Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1998;96(1):31–6.

Wang T, Zhao Z, Yan H, et al. Gene flow from cultivated herbicide-resistant foxtail millet to its wild relatives: a basis for risk assessment of the release of transgenic millet. Acta Agron Sin. 2001;27(6):681–7.

Wang C, Chen J, Zhi H, et al. Population genetics of foxtail millet and its wild ancestor. BMC Genet. 2010;11(1):90.

Wang C, Jia G, Zhi H, et al. Genetic diversity and population structure of Chinese foxtail millet [*Setaria italica* (L.) Beauv.] landraces. G3 Genes Genom Genet. 2012;2(7):769–77.

Watt G. A Dictionary of the Economic Products of India (vols I–VI). Commercial Products of India: Calcutta; 1908.

Wei S. The origin, domestication and spreading of Chinese foxtail millet. Anc Curr Agric. 1994;2:6–17.

Willweber-Kishimoto E. Interspecific relationships in the genus Setaria. Contrib Biol Lab Kyoto Univ. 1962;14:1–41.

Wu WW, Wang XH, Wu XH, et al. The early Holocene archaeobotanical record from the Zhangmatun site situated at the northern edge of the Shandong Highlands, China. Quat Int. 2014;348:183–93.

Yang X, Wan Z, Perry L, et al. Early millet use in northern China. Proc Natl Acad Sci. 2012;109(10):3726–30.

Zhang G, Liu X, Quan Z, et al. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol. 2012;30(6):549–54.

Zhao Z. The process of origin of agriculture in China: archaeological evidence from flotation. Quat Sci. 2014;34:73–84.

# Chapter 5
# Foxtail Millet Germplasm and Inheritance of Morphological Characteristics

**Xianmin Diao and Guanqing Jia**

**Abstract** In China, the systematic collection of foxtail millet germplasm from the 1950s to the 1980s resulted in the compilation of 27,059 accessions in the Chinese Gene Bank. There are approximately 15,000 additional foxtail millet accessions maintained in other gene banks in India, Japan, Korea, the United States of America, Russia, and in other countries. Evaluations of the Chinese and Indian accessions indicate that foxtail millet is morphologically and genetically highly diverse, especially in China. There are currently only two foxtail millet core collections, in China and India. Large-scale screenings of trait-specific lines have been conducted mainly in China, and have identified some special landraces, including those that are resistant to drought. There are many publications describing the inheritance of foxtail millet morphological characteristics in China, with the important ones being reviewed in this chapter. Dominant qualitative traits, including seedling, leaf sheath, and anther color, have been used as markers to identify hybrids. The estimated heritabilities of quantitative characteristics, such as plant height, panicle length, and heading date, have been useful for foxtail millet breeding programs. Additionally, the recent detection of significant quantitative trait loci has helped to characterize the underlying genetic mechanisms regulating specific traits. Cytological studies of foxtail millet are summarized in this chapter, especially those involving the trisomic Yugu 1 cultivar.

**Keywords** Foxtail millet • Germplasm management • Qualitative characteristics • Quantitative characteristics • Trisomics • Setaria

## 5.1 Introduction

Because of its long cultivation history and considerable contribution to ancient civilizations in China, foxtail millet is considered one of the most important grains among the “Five Grains of China” (Diao 2011; Austin 2006), along with

X. Diao (✉) • G. Jia
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences,
No. 12, Zhongguancun South St., Haidian District, Beijing 100081,
People’s Republic of China
e-mail: diaoxianmin@caas.cn



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_5

proso millet, rice, soybean, and wheat. Foxtail millet is used even today in ancestor worship ceremonies and is considered a sacred crop in some communities and regions. A long history of cultivation in diverse environments around the world has resulted in numerous foxtail millet cultivars, which also serve as useful genetic resources. Foxtail millet cultivation in the dry areas of the loess-rich upper Yellow River Basin and in the semidry areas of the Chinese middle basin led to the development of cultivars with high levels of drought resistance and cold tolerance. These cultivars also produce relatively large seeds and exhibit improved lodging resistance. The importance of foxtail millet with desirable cooking and tasting qualities resulted in the development of nonglutinous and glutinous types, as well as varieties that produced diverse kernel colors and various aromas upon cooking. Some varieties with good cooking qualities and a pleasant taste became popular in certain regions, including a glutinous variety used to produce millet wine, which is described in the Shang dynasty oracle bones (Bonjean 2010). For breeding and genetic research purposes, the inheritance of many morphological characteristics, including qualitative and quantitative traits, has been characterized in studies by different researchers.

## 5.2 Foxtail Millet Germplasm and Classification

### *5.2.1 Collection of Foxtail Millet Germplasm*

There is a long history of foxtail millet germplasm collection, including landraces, improved varieties, and wild relatives. The Guangzhi (i.e., book related to agriculture) was written during the Jin Dynasty (1700 years ago), and described 11 foxtail millet varieties for the first time in China. The famous Qiminyaoshu, which was written during the Beiwei Dynasty, described 64 foxtail millet varieties, and also classified these varieties into early heading, drought-resistant, bird-resistant, and good-tasting categories (Diao 2011). During the Qing Dynasty, the Shoushitongkao described 251 foxtail millet varieties, which were grown in different regions. Since the 1920s, the University of Nanking and Yenching University have collected and identified several foxtail millet landraces for breeding programs.

The collection of foxtail millet germplasm in China began in the mid-1950s, and by 1958, about 16,000 accessions had been accumulated. Because of the Cultural Revolution, the germplasm collection and related research were not updated during the 1960s and early 1970s. This work resumed in 1987, and 11,673 foxtail millet accessions have since been assigned an ID number and catalogued in three books (List of Chinese Foxtail Millet Varieties, issues 1, 2, and 3). By the end of 2012, there were 27,059 foxtail millet accessions in the Chinese Gene Bank, and seven issues of the List of Chinese Foxtail Millet Varieties had been published. This Chinese collection is a nationwide collection that covers all

foxtail millet growing regions in China. Additionally, most accessions collected before 2000 were landraces, while some improved cultivars were also collected after 2000. The number of accessions in gene banks continues to grow.

There is no comprehensive publication describing global foxtail millet germplasm collections. However, additional relatively small collections are maintained at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India (1535 accessions from 26 countries), the National Institute of Agrobiological Sciences, Tsukuba, Japan (1286 accessions), the Plant Genetic Resources Conservation Unit of the United States Department of Agriculture-Agricultural Research Service (766 accessions), the National Gene Bank of Korea (about 960 accessions), and the National Gene Banks of India (2774 accessions) and Bangladesh (510 accessions). Foxtail millet germplasm is also available elsewhere, including in gene banks in Russia and other European countries, but the number of accessions is unknown.

### *5.2.2 Classification of Foxtail Millet Collections*

Because of natural and man-made selection pressures in various environments, several varieties have evolved from foxtail millet accessions. Dekaprelerich and Kasparian (1928) classified foxtail millet into the following two groups: Moharium (i.e., small panicles and many tillers/branches) and Maxima (i.e., large panicles and single/fewer tillers/branches). Rao et al. (1987) later defined three races of foxtail millet varieties and suggested Indica (i.e., intermediate between Moharium and Maxima) should be considered a third group for cultivar classifications. These studies formed the basis of foxtail millet germplasm collection and breeding efforts. Details regarding the three categories are provided as follows.

1. Maxima varieties contain large panicles with small inflorescences compactly spread along the primary branches. Each individual contains 1–8 tillers (one panicle per tiller) with no branches. The panicles are either large and sagging with long bristles (i.e., varieties originating from northeast China, Japan, and North Korea) or small and upstanding with short bristles (i.e., varieties originating from the Inner Mongolia Autonomous Region of China). Maxima varieties have been widely grown as food in eastern Europe and eastern Asia, and as a forage crop or bird feed in North America.
2. Moharia varieties produce small panicles with small inflorescences compactly spread along the main branches. They also contain long bristles. Morphologically, these varieties are similar to green foxtail, but the inflorescences do not shatter, as in other domesticated varieties. Each individual contains 5–52 tillers with branches. Erect panicles are present on each branch. Moharia varieties have been cultivated in southeastern Europe, Russia, India, Pakistan, and Afghanistan.
3. Indica varieties are intermediates that exhibit Maxima and Moharia characteristics. They originated from hybridization between Moharia (from India) and Maxima

(from China) varieties. Each individual contains 1–12 tillers, and moderately upstanding panicles (with long bristles) are present on tillers. Primary branches are scattered along the main branch. Indica varieties are mainly grown in India.

The three categories are only meant as a guide, and some accessions are difficult to classify because of the abundance of intermediate foxtail millet types. Li et al. (1995) suggested foxtail millet germplasm should be classified into four races (i.e., maxima, moharia, indica, and nana) based on morphological analyses of 2907 accessions from a global collection. The nana race includes accessions from Lebanon and Afghanistan, with intermediate primary characteristics between the wild and domesticated types. Breeding programs have used several methods to classify foxtail millet germplasm, including classification according to tillering (nontillering or tillering) (Ochiai 1996), panicle type (spindle, cylinder, stick, finger, cat foot, or hen beak), seed color (white, yellow, orange, red, or black), kernel color (yellow or white), and phenol color reactions (Kawase and Sakamoto 1982). Other popular classification methods involve multivariate morphological analyses and molecular markers. Details relevant to the classification of foxtail millet are provided in the following sections, as well as in Chaps 2 and 7.

### *5.2.3 Germplasm Diversity and Construction of Core Collections*

Foxtail millet is morphologically quite diverse. A clustering analysis of 2907 accessions from 16 Chinese provinces and 22 other countries using nine morphological and agronomic traits indicated all characteristics were highly variable. Additionally, the samples were grouped into 12 clusters based on differences in geographical origins (Li et al. 1995). Considerable diversity was also observed for 15 morphological and agronomic characteristics in 878 global foxtail millet accessions (Wang et al. 2016). Of the nine quantitative traits, grain weight per panicle and panicle length exhibited the most variability and the highest diversity indices (Table 5.1).

Genetic diversity of foxtail millet has been assessed using molecular markers in experiments involving restriction fragment length polymorphisms (RFLPs) (Schontz and Rether 1998; Fukunaga et al. 2002a, b), random amplified polymorphic DNA (Li et al. 1998; Schontz and Rether 1999; Kumari et al. 2011), intersimple sequence repeats (Li et al. 2012; Kumari et al. 2011), amplified fragment length polymorphisms (d'Ennequin et al. 2000), simple sequence repeats (Wang et al. 2012; Jia et al. 2015), single nucleotide polymorphisms (SNPs) (Wang et al. 2010; Jia et al. 2013; He et al. 2015), rDNA PCR–RFLP (Eda et al. 2013), ribosomal DNA variation (Fukunaga et al. 1997, 2005, 2006, 2011), transposon display (Hirano et al. 2011), and other molecular markers (Liu et al. 2011, 2014). Although the sample sizes, objectives, and molecular markers used

**Table 5.1** Variations and distribution frequencies of quantitative traits in 878 foxtail millet accessions (Wang et al. 2016)

| Agronomic trait | Mean ± SD | Range | CV (%)[a] | $H'$[b] |
|---|---|---|---|---|
| Plant height (cm) | 139.17 ± 20.94 | 45.00–195.00 | 15.04 | 1.73 |
| Diameter of main stem (cm) | 0.68 ± 0.12 | 0.25–1.20 | 17.71 | 1.59 |
| Stem node number | 11.73 ± 1.84 | 4.00–17.00 | 15.71 | 1.74 |
| Peduncle length (cm) | 10.84 ± 4.47 | 1.00–36.25 | 41.24 | 1.03 |
| Panicle length (cm) | 23.98 ± 5.62 | 5.25–43.00 | 23.45 | 1.80 |
| Panicle diameter (cm) | 2.54 ± 0.61 | 0.95–5.15 | 23.95 | 0.67 |
| Panicle weight per main stem (g) | 15.03 ± 5.95 | 0.20–39.50 | 39.58 | 1.84 |
| Grain weight per main stem (g) | 11.85 ± 5.14 | 0.15–34.20 | 43.41 | 1.82 |
| Growth period (d) | 121.79 ± 9.67 | 56.00–155.00 | 7.94 | 1.39 |

[a]Coefficient of variation (= standard deviation divided by the mean) × 100 %
[b]Shannon-Wiener diversity index ($H'$)

differed among studies, some common conclusions can be drawn. First, foxtail millet germplasm is highly diverse, as confirmed by a sequence diversity ($\pi$) of approximately 0.0010 for all samples. This result is intermediate between the values for the cultivated rice species *Oryza sativa* ssp. *indica* (approximately 0.0016) and *O. sativa* ssp. *japonica* (approximately 0.0006) (Jia et al. 2013). Second, most of these studies revealed that Chinese samples were the most diverse, implying China is the center of diversity for foxtail millet landraces. Third, most of these studies identified clear geographic differences, indicating that environmental factors and local cultivar selection by humans considerably influenced genetic evolution.

To date, two foxtail millet germplasm core collections have been constructed, one at the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, China (Jia et al. 2013; Wang et al. 2016), and the other (containing ICRISAT samples) maintained in India (Upadhyaya et al. 2009). The Chinese core collection consists of 916 samples, including Chinese landraces, modern cultivars, and accessions from other countries. The samples were selected from 27,000 accessions in the Chinese Gene Bank. Low coverage sequences of all 916 samples are available online, and a haplotype map of the foxtail millet genome consisting of 0.8 million common SNPs has been constructed (Jia et al. 2013). The ICRISAT gene bank includes 1535 foxtail millet accessions from 26 countries. Using passport information and data regarding 23 morphological descriptors, Upadhyaya et al. (2009, 2011) developed a core collection consisting of 155 foxtail millet accessions. Most of the accessions in this collection are from the indica race (i.e., 102 accessions; 65.8 %). The next most abundant races were maxima (i.e., 24 accessions; 15.5 %) and moharia (29 accessions; 18.7 %). However, there is currently no available genomic data regarding these core accessions. The Chinese and ICRISAT core collections represent the genetic diversity and population structure of foxtail millet germplasm from around the world.

**Table 5.2** Identification of Chinese accessions with specific quality traits (Lu 2016)

| Items | Number of accessions tested | Range | Standards for super quality | Number of super accessions |
|---|---|---|---|---|
| Rude protein | 21,076 | 7.9–21.9 % | >15 % | 2887 |
| Rude fat | 21,076 | 1.14–7.99 % | >5 % | 2656 |
| Lysine contain | 20,876 | 0.11–0.46 % | >0.33 % | 1124 |
| Starch contain | 939 | 64.64–85.95 % | >80 % | 188 |
| Vitamin $B_1$ | 881 | 2.7–10.2 mg/kg | >8 mg/kg | 46 |
| Vitamin $B_2$ | 881 | 0.48–1.76 mg/kg | >1.2 mg/kg | 28 |
| Vitamin E | 1081 | 4.25–95.85 mg/kg | >65 mg/kg | 144 |
| Selenium | 1081 | 13–212 mg/kg | >100 mg/kg | 209 |

### *5.2.4 Identification of Trait-Specific Germplasm*

Identifying trait-specific accessions from the large number of accessions in the available germplasm collections is crucial for efficient breeding and genetics research. Many accessions of foxtail millet varieties have been screened for grain quality, resistance to diseases, pests, and drought, and other agronomic traits. Some of the relevant results are summarized in Tables 5.2 and 5.3.

For nutritional-related characters, a subset of selected accessions was tested, and ranges for crude protein of 7.25–21.9 %; crude fat 1.14–7.99 %; starch 64.64–85.95 %; vitamin B1 2.7–10.2 mg/kg; vitamin B2 0.48–1.76 mg/kg; vitamin E 4.25–95.85 mg/kg; and Se (minor mineral) 13–212 mg/kg (Lu 2016) were found. Accessions with superior agronomic (e.g., earliness and high grain yield) and nutritional (e.g., high seed protein, calcium, iron, and zinc contents) traits were identified by ICRISAT, and a diverse range of foxtail varieties have been selected for breeding. Details regarding these accessions have been published (Upadhyaya et al. 2011).

Since the 1950s, specific screening tests have been used to identify accessions that are resistant/sensitive to various foxtail millet diseases and pests (Lu 2006), including the spraying of seedlings with spores responsible for blast and leaf rust diseases, infection of germinating seedlings using soil inoculated with spores causing panicle smut, and the natural infection of foxtail millet growing in a nursery with *Aphelenchoides besseyi* in regions where this pathogen is commonly detected. All of these screening tests were conducted using three replicates and repeated in 2 years. In addition to the extensive screening of the older germplasm in the Chinese Gene Bank for tolerance to diseases and drought, many studies have been conducted involving recently released cultivars. Examples of such studies include the analyses of resistance to *Thanatephorus cucumeris* (Frank) Donk (Ma et al. 2005), leaf rust (Dong et al. 2012), blast disease (Dong et al. 2015), nematodes (Dong et al. 2010), and *A. besseyi* (Cui

**Table 5.3** Identification of disease- and drought-resistant foxtail millet accessions

| Disease/drought | Method used | Number of accessions tested | Number of highly resistant/tolerant accessions | Number of resistant/tolerant accessions |
|---|---|---|---|---|
| Blast | Spore spray | 18,470 | 166 | 691 |
| *Sclerospora graminicola* | Mixed soil infection | 22,797 | 551 | 834 |
| Panicle smut | Mixed soil infection | 6031 | 44 | 61 |
| Leaf rust | Spore spray | 12,021 | 74 | 17 |
| *Aphelenchoides besseyi* | Natural infection | 10,181 | 1 | 28 |
| Corn borer | Artificial infection | 3072 | 0 | 6 |
| Nematode | Natural infection | 1050 | 69 | 108 |
| Drought tolerant at seedling stage | Repeated stress at seedling stage | 21,470 | 807 | 2356 |
| Drought tolerant during growth duration | Controlled irrigation in the field | 207 | 2 | 47 |

et al. 1989). Many cultivars susceptible and resistant to various diseases and pests have been identified, which has benefited the foxtail millet industry (Cui 1997; Gan 1997).

Drought tolerance is an important consideration for many plant breeders. More than 21,470 basic foxtail millet accessions from germplasm collections have been screened to identify genotypes tolerant or sensitive to drought conditions (Li 1991, 1992). Some of these results are summarized in Table 5.3. Additionally, commercially grown cultivars have also been tested for tolerance to drought and salinity stresses (Zhi et al. 2004). To establish an effective method to screen for lines tolerant to drought for the duration of the growing period, related indices were analyzed in the identified tolerant and sensitive cultivars. The booting stage was revealed to be the most sensitive period for foxtail millet reproduction (Zhang et al. 2010). Some indices were developed to identify foxtail millet accessions that were drought tolerant during the booting stage (Zhang et al. 2012). The effects of drought stress on foxtail millet photosynthetic characteristics and other physiological traits at the booting stage were also tested in drought-tolerant and -sensitive cultivars (Zhang et al. 2011a, b).

While several studies have indicated the value of germplasm collections for understanding agronomic traits, Setaria is also being developing as a novel model system. Germplasm accessions are essential not only for association mapping but also for functional gene evolution studies. However, more cooperation is needed between countries to increase sampling across the range of variation.

## 5.3 Inheritance of Morphological Characteristics in Foxtail Millet

Because foxtail millet is agronomically important only in certain regions, research on inheritance of morphological characters in foxtail millet has mostly been published in local Chinese journals. Although there are many scientific papers on foxtail millet in these journals, we have selected only those we think important for review. Although some of the reported results are preliminary or even disagree with each other, we think they provide valuable information for future investigations into the genetic basis of these characters.

### *5.3.1 Qualitative Characteristics*

#### 5.3.1.1 Seedling and Leaf Sheath Colors

Foxtail millet seedling color is controlled by four Mendelian factors (i.e., PPVVHHII), with PP corresponding to the gene for purple seedlings, II enhances the color, VV ensures only the stem sheath turns purple, and HH ensures only the panicle becomes purple (Ayyangar et al. 1935). In most seedlings, the blades of lines with purple leaf sheaths are green, and the purple leaf sheath is dominant over the green leaf sheath. This trait is frequently used as a marker to identify true hybrid plants in crosses using a green female parent and a purple male parent.

Association mapping using 916 foxtail millet accessions from global core collections detected two major quantitative trait loci (QTLs) controlling leaf sheath color located at genome positions 26,859,601 and 7,390,404 on chromosomes 7 and 4, respectively. For seedling leaf blade color, one major QTL on chromosome 7 and several QTLs with small effects were identified in the three tested environments (Jia et al. 2013).

#### 5.3.1.2 Grain Color, Kernel Color, and Waxy Endosperm

There is a wide range of foxtail millet seed colors, including white, brown, light yellow, deep yellow, red, gray, and black. Foxtail millet seed color is controlled by three genetic loci (BBIIKK), with BB responsible for gray seeds, II enhances the color, and KK causes seeds to become yellow (Ayyangar and Narayanan 1931). Additionally, Zhang (1961) reported that $F_1$ plants derived from the Shilixiang (red seeds) × Dabaigu (light yellow seeds) cross produce seeds with an intermediate seed color. Researchers from the Inner Mongolia Agricultural Research Institute (1979) determined that in foxtail millet, light yellow seed color is dominant over red seed color, and white seed color is dominant over yellow seed color. There have only been a few studies examining foxtail millet grain color, and the results suggested that seed color is controlled by several major genes and regulatory elements. Two

major QTLs regulating seed color were identified on chromosomes 1 and 9 in three tested environmental conditions. An additional five QTLs with minor effects were detected on chromosomes 6 and 7 (Jia et al. 2013).

Foxtail millet kernel color was analyzed by Li et al. (1940), who determined that yellow is dominant over white. However, there is a wide range in the color intensity of yellow among different varieties. Additionally, foxtail millet kernel color is regulated by two gene loci (i.e., Y1 and Y2) (Takahashi 1942) and is of great interest to breeders and farmers. The yellow appearance of foxtail millet kernels is mainly due to carotenoid-related chemicals. The wild green foxtail kernel is white, with no or very little carotenoid. This suggests that the yellow kernel of foxtail millet is a domestication-related characteristic.

Waxy endosperm arises through the disrupted expression or loss of function of the waxy (*GBSS1*) gene, which encodes granule-bound starch synthase I. There are different types of waxy endosperm in foxtail millet (Afzal et al. 1996), and the waxy trait is usually dominant over the nonwaxy trait. Mao et al. (2000a, b) mapped the waxy gene to chromosome 4, using a primary trisomic series of foxtail millet cultivar Yugu 1 and SDS-PAGE electrophoresis analysis. Genomic sequence analysis of the *Wx* genes revealed five independent origins for different endosperm types. Additional details regarding foxtail millet endosperm differentiation is provided by Fukunaga et al. (2002a, b) and in Chap. 7.

#### 5.3.1.3 Bristle Length and Color

According to a study by Lupeng Zhang (1972), the production of long bristles in foxtail millet is regulated by dominant alleles. This conclusion was based on the fact that all $F_1$ hybrids derived from a cross between parents with long and short bristles produced long bristles. Ayyangar and Narayanan (1933) reported that bristle length is determined by four genes. Foxtail millet bristle color was analyzed by Lupeng Zhang (1972), who concluded that purple bristles were encoded by a dominant allele, while green bristles were the consequence of a recessive allele. Bristle length is also a domestication-related characteristic. The bristles of wild green foxtail are relatively long and densely distributed. In contrast, bristle length and density of cultivated foxtail millet are highly variable among landraces. Farmers used to prefer landraces with short bristles. One major QTL for bristle color was identified on chromosome 4 under five environmental conditions. Additionally, one major QTL for bristle length was detected on chromosome 1 in most of the tested environmental conditions (Jia et al. 2013).

#### 5.3.1.4 Anther Color

Foxtail millet anthers can be brownish-orange or white. These colors form a simple Mendelian pair, with orange being the dominant trait (Ayyangar and Narayanan 1932). Anther color must be observed in the morning because after pollen

shattering, all the white and orange anthers turn brown. Anther color has been used as a marker for identifying individual hybrids from crosses between a female parent with white anthers and a male parent with orange anthers. There are also purple anthers among the *Setaria* species (e.g., *Setaria glauca*). Association mapping revealed that a gene regulating anther color is located at position 34,378,428 on foxtail millet chromosome 6 (Jia et al. 2013).

#### 5.3.1.5 Panicle Type

There are several types of foxtail millet panicles, including cone-like, spindle-like, long spike-like, palmate-like, cylindrical, highly branched (finger), rod-like, cat foot, and hen beak. The panicle type can be greatly affected by environmental conditions. Li et al. (1935) observed that palmate-like panicles are controlled by two dominant genes. Zhang (1961) characterized the $F_2$ offspring derived from parents with different panicle types, and concluded that panicle type in foxtail millet is regulated by multiple genes. There are more than ten foxtail millet panicle types in the Chinese Gene Bank collection, but in none of them are the inheritance patterns clearly understood. Breeding experiments revealed that most panicle types are controlled by major genes, with complex inheritance patterns. More detailed research is needed to clarify their inheritance.

#### 5.3.1.6 Dwarf Accessions

Nearly 50 dwarf lines have been identified in Chinese foxtail millet breeding programs (Qian et al. 2012). A longtime objective of foxtail millet breeders has involved the identification of suitable genes that could be used to breed lodging-resistant dwarf cultivars. Phenotypic evaluation and breeding experiments indicated that most dwarf foxtail millet lines exhibit premature senescence, which limits their utility for breeding. The dwarf trait is most often controlled by major recessive genes, such as *Anai3* and *Ai25* (Gao et al. 2003; Xue et al. 2016). Gibberellin (GA) sensitivity tests indicated that all dwarf foxtail millet lines are GA sensitive, except for line 84,133 from the Agricultural Institute of Chifeng in Inner Mongolia (Chen et al. 1998; Qian et al. 2012; Yao and Liang 1990). The gene encoding the GA insensitivity of line 84,133 was cloned and determined to encode a DELLA protein (Zhao et al. 2016). The *Anai3* gene of a foxtail millet dwarf line was fine mapped onto chromosome 3 (Gao et al. 2003). Additionally, the *Ai25* gene was fine mapped to a 52.7-kb region between markers fxj032 and fxj037 on chromosome 3. The inheritance and genetic background of most other foxtail millet dwarf lines have yet to be fully characterized. Dwarf genotypes have also been identified in Indian foxtail millet breeding programs, and the agricultural performance of some of these lines has been analyzed (Dineshkumar et al. 1992).

### 5.3.2 *Quantitative Characteristics*

Most agronomic traits are regulated by multiple genes or QTLs, including plant height, panicle length, heading date, and nutrition-related characteristics. Compared with the available publications related to rice and other major crops, there are relatively few reports regarding the inheritance of these characteristics in foxtail millet. Additionally, most of the studies that generated relevant data in China were conducted in the latter part of the previous century. The results of most of these investigations were published in local journals, such as *Acta Agronomica Sinica*, *Acta Agriculturae Boreali-Sinica*, *Genetic Resources and Crop Evolution*, *Hebei Agricultural Sciences*, *Shanxi Agricultural Sciences*, and others. Li Renzhi, Zhao Zhili, and Mao Xue summarized some of these studies in a book which entitled "Principles of foxtail millet genetics and breeding."

#### 5.3.2.1 Plant Height

Many studies on foxtail millet plant height have concluded that it is controlled by multiple genes or QTLs. According to Liu (1984), the heritability of plant height in foxtail millet is over 80 %. The $F_2$ segregation results for foxtail millet plant height generated in 1979 by the Foxtail Millet Laboratory of the Crop Sciences Institute, Inner Mongolia. The plant height of $F_1$ hybrids is usually in between that of the two parents, with tall plants usually being the dominant. The segregation in $F_2$ plants varies according to their parents. The segregation for a cross between short and tall parents is usually greater than that of a cross between two tall parents. Using genome-wide association mapping, four significant QTLs regulating plant height in foxtail millet were identified on chromosomes 2, 4, 6, and 8 under different environmental conditions (Jia et al. 2013). A QTL analysis of a biparental mapping population derived from a cross between foxtail millet and green millet revealed QTL on chromosomes 1, 2, 3, 4, 5, 6, 7, 8, and 9 in two different environments (Mauro-Herrera and Doust 2016).

#### 5.3.2.2 Panicle Length

According to field observations reported by the Institute of Crop Sciences, Inner Mongolia (1979), panicle length in $F_1$ plants is between the lengths of the parents, but tends to be similar to the length of one of the parents in the hybrids of some crosses. Heterosis was also detected for foxtail millet panicle length. The heritability of panicle length is approximately 75 % (Liu 1984; Laboratory of Foxtail Millet, Institute of Crop Sciences, Inner Mongolia 1979).

Panicle length is a domestication-related characteristic that is also closely related with grain yield, which makes it of great interest to breeders. Genome-wide association mapping identified only one QTL for panicle length for foxtail millet grown in Beijing

and Changzhi, but three and ten QTLs for plants grown in Anyang and Sanya, respectively. These results imply that the QTLs are specific to certain environments, and that foxtail millet panicle length is influenced by environmental conditions (Jia et al. 2013).

#### 5.3.2.3 Heading Date

Heading date is obviously a domestication-related characteristic. Wild green foxtail is photoperiod sensitive and exhibits early heading. It has a short growing period resulting in low biomass. Natural mutations and human selections have generated late heading lines with a relatively long growing period that ultimately leads to higher biomass. The spread of foxtail millet north and south from its original domestication site required genetic changes that affected photoperiod and temperature sensitivity and heading date. Foxtail millet landraces from the north are usually more photoperiod sensitive and exhibit early heading with a short growing period. Additionally, in most cases, early heading is dominant over late heading.

According to studies conducted at the Inner Mongolia Agricultural Research Institute (1979), the heading dates of $F_1$ hybrids are determined by the parents. The correlation coefficient between the heading dates for the $F_1$ plants and the average heading dates for the parents is 0.787. Additionally, there are significant relationships between heading dates for $F_1$ plants and the parent ecotypes. If the ecotypes of the two parental landraces are identical or similar, their $F_1$ hybrids will have an intermediate heading date. If there are extensive differences in the parental ecotypes, the early heading trait will usually exhibit a semidominant inheritance pattern. Transgressive inheritance of heading date was also observed in the $F_2$ offspring derived from parents originating from different eco-regions. The heritability of heading date in foxtail millet may be as high as 94.66%, based on breeding experiments conducted by Liu (1984).

Linkage analysis of recombinant inbred lines generated from crosses between foxtail millet and green foxtail identified 16 heading date-related QTLs distributed among all nine chromosomes. Many of these QTLs likely have syntenic orthologs in sorghum and maize (Mauro-Herrera et al. 2013). Association mapping with 916 foxtail millet genotypes in 2011 identified six, seven, seven, and 20 heading date-related QTLs in plants cultivated in Beijing, Anyang, Changzhi, and Sanya, respectively. The differences in the number of detected QTLs suggest that many of these QTLs are specific to particular environments (Jia et al. 2013).

#### 5.3.2.4 Tillering

Tillering is a domestication- and grain yield-related characteristic. Wild green foxtail often has many tillers, but the tillering in domesticated foxtail millet is ecotype dependent. Ecotypes that have adapted to dry and barren land often have many tillers (e.g., landraces from central Asia). In contrast, ecotypes grown on well-managed fertile land often have few or no tillers (e.g., landraces from the North China Plain). Breeding experiments have indicated that tillering is somewhat dominant over nontillering, depending on the parent genotypes.

Linkage analysis using a mapping population derived from a cross between foxtail millet B100 and green foxtail A10 identified four significant QTLs associated with tillering in two experimental trials. Comparative genomics investigations of maize indicated that an ortholog of *teosinte branched1*, which is the major gene regulating tillering in maize, has only a minor and variable effect in foxtail millet. This implies that similar phenotypic effects may not be produced by orthologous loci, even in closely related species (Doust et al. 2004). Additional growth trials with recombinant inbred lines derived from this population in field and green house environments have identified further QTL, supporting the variable nature of tillering under differing environments (Mauro-Herrera and Doust 2016). This was confirmed by association mapping using foxtail millet accessions from core collections that identified many QTLs with small effects, with only a few that colocalized, indicating tillering is considerably influenced by environmental conditions (Jia et al. 2013). The fact that tillering is controlled by QTLs with small effects is ideal for breeding. By using different combinations of QTLs, breeders can develop varieties with different degrees of tillering that suitable for specific environments or fields treated with various field-management practices.

#### 5.3.2.5 Nutrition-Related Traits

An investigation of the heredity of total protein content in 15 hybrid combinations revealed that the $F_1$ hybrids exhibited maternal inheritance patterns (Gu et al. 1992). The segregation ratio for protein content in $F_2$ offspring suggested that protein content is controlled by multiple loci with minor effects, and the average broad sense heritability of total protein content is 50 %. Heterosis over the better parent (23.1 %) regarding protein content was also observed in all the 15 hybrids, indicating that hybrid vigor for protein content can be applied to improve foxtail millet varieties.

Results of an examination of the lipid content of nine hybrids suggested that $F_1$ hybrids mostly exhibit mid-parent values, and lipid content in foxtail millet is regulated by multiple loci (Gu et al. 1991). Broad sense heritability of lipid content can be as high as 81 %, which implies that selection of high-oil foxtail millet varieties in early generations may be possible in breeding programs.

### *5.3.3 Heritability and Correlations Among Quantitative Traits*

Based on field performance assessments of Chinese foxtail millet accessions (Zhao and Jin 1985; Laboratory of Millet Crops, Institute of Crops of Hebei Province 1975; Liu 1984), the broad sense heritability of primary agronomic traits in foxtail millet grown in different eco-regions of China has been determined (Table 5.4). Desirable traits exhibiting relatively high heritability in foxtail millet can be selected in early generations in pedigree breeding programs.

**Table 5.4** Heritability of primary agronomic traits in foxtail millet (Zhao and Jin 1985; Laboratory of Millet Crops, Institute of Crops of Hebei Province 1975; Liu 1984)

| | Eco-regions of China | | |
|---|---|---|---|
| Rank | Northeast | Summer sowing | Spring sowing |
| 1 | Plant height (85.39 %) | Grain weight per hectare (94.8 %) | Duration of growth period (98.15 %) |
| 2 | Duration of growth period (84.12 %) | Number of primary branches per panicle (89.9 %) | Heading date (94.66 %) |
| 3 | Days from heading to mature (82.18 %) | Plant height (83.4 %) | Plant height (88.15 %) |
| 4 | Heading date (81.42 %) | 1000-seed weight (78.4 %) | Days from heading to mature (86.66 %) |
| 5 | Node number (76.19 %) | Panicle length (77.4 %) | Node number (85.35 %) |
| 6 | Stem diameter (73.23 %) | Root number (61.0 %) | Number of primary branches per panicle (81.12 %) |
| 7 | Panicle length (72.08 %) | Number of layer of air root (59.0 %) | Panicle diameter (77.87 %) |
| 8 | Weight of stalk (71.33 %) | Grain weight per plant (47.3 %) | Number of layer of air root (76.74 %) |
| 9 | 1000-seed weight (58.3 %) | Panicle weight per plant (39.8 %) | Panicle length (74.18 %) |
| 10 | Panicle weight (52.58 %) | | Grain weight (74.12 %) |
| 11 | Grain weight per panicle (50.11 %) | | Panicle weight (73.96 %) |
| 12 | Grain yield per plant (41.62 %) | | 1000-seed weight (68.42 %) |
| 13 | | | Weight of stalk (65.75 %) |
| 14 | | | Stem diameter (61.56 %) |

Correlations among important agronomic characteristics have been analyzed by Chinese breeders. Duan et al. (1990) observed that plant height was positively correlated with panicle length, number of primary branches per panicle, and number of grains per panicle. Panicle length was positively correlated with the number of primary branches per panicle, grain weight per panicle, and 1000-seed weight. The number of primary branches per panicle was positively correlated with grain weight per panicle, while the 1000-seed weight was positively correlated with grain weight per panicle. Liu et al. (1990) reported that the grain yield per plant was positively correlated with the number of grains per panicle, grain weight per panicle, and 1000-seed weight. However, the number of grains per panicle was negatively correlated with 1000-seed weight. Li and Ling (1989) determined that the grain protein content was positively correlated with 1000-seed weight but negatively correlated with grain yield per plant. Most of those findings are preliminary and but have provided useful information for breeders working in foxtail millet improvement programs in China.

## 5.4 Cytogenetics of Foxtail Millet

There are typically nine chromosomes in the nucleus of monoploid *Setaria* species (i.e., gametophytic egg and pollen; n = x = 9) (Avdulov 1931). Green foxtail and foxtail millet share the same A genome and are diploid species (2n = 2x = 18) (Li et al. 1945; Wang et al. 2009; Zhao et al. 2013).

### *5.4.1 Foxtail Millet and Green Foxtail Karyotypes*

A karyotype presents the morphological characteristics of chromosomes in a species, including chromosome number, morphology (i.e., length, diameter, size, and satellite position), structure, and Giemsa banding patterns. Karyotypes for foxtail millet and green foxtail have been described (Sun et al. 1983; Li and Chen 1985). Based on analyses of 23 foxtail millet varieties, the total length of the nine chromosomes is 21.92–43.93 μm, while the individual chromosomes are 3.59–5.74 μm. Additionally, only chromosome 7 contains a satellite. The haploid karyotype of foxtail millet may consist of n = 9 = 7 m + 1st + 1st satellite. The karyotype of green foxtail is identical to that of foxtail millet, which is consistent with the fact that green foxtail is the wild ancestor of foxtail millet. The observed chromosome length may be influenced by the chemicals and pretreatments used before analyses. Therefore, the relative chromosome length (i.e., proportion of the total chromosome length) is widely used for karyotype studies. The lengths of foxtail millet chromosomes 1 and 2 correspond to 15–16% and 13–14% of the total chromosome length, respectively. There is a gradual decrease in the lengths of chromosomes 3 to 7, and chromosome 8 is the shortest. Chromosome 9 is slightly longer, (9–10% of the total chromosome length) (Sun et al. 1983). The relative lengths of green foxtail chromosomes 1–9 are 16.54 ± 0.65%, 13.20 ± 0.76%, 12.50 ± 0.08%, 11.74 ± 0.48%, 10.23 ± 0.06%, 9.42 ± 0.58%, 8.76 ± 0.72%, 8.27 ± 1.51%, and 9.33 ± 1.73%, respectively. These values are nearly identical to the corresponding values for foxtail millet. A comparison between the karyotypes of Chinese and European accessions revealed differences between samples from diverse origins (Sun et al. 1994).

### *5.4.2 Primary Trisomic Systems of Foxtail Millet and Their Utility*

Primary trisomic systems are genetic tools that have been used to locate genes and identify genetic linkages, especially before molecular markers were developed. A foxtail millet primary trisomic system was constructed through the hybridization of an artificially 4n Yugu 1 with a naturally 2n Yugu 1 in 1994. This system consisted of nine lines corresponding to nine foxtail millet chromosomes (Wang et al. 1994, 1999; Gao et al. 2000). Individual trisomic lines carry three copies of one of its

**Table 5.5** Characteristics of nine foxtail millet trisomic lines

| Trisomics | Diagnostic characteristics |
|---|---|
| Trisomic I (curl leaves) | Narrow and short leaves, most leaves curl and erective except basal ones, dense and small panicle with some degree of sterile, no tillers |
| Trisomic II (dark green) | Narrow, dark green, short and drooping leaves at seedling, short spindle panicle and ellipse grain, some individual have tillers |
| Trisomic III (tiller type) | The shortest plants of the series with many tillers, yellow green and dropping leaves, many bristles on the top of the panicle and some degree of sterile |
| Trisomic IV (long bricles) | Tall plant with tillers, narrow, short and drooping leaves, loose panicle with long bristles and nearly normal seed setting |
| Trisomic V (slim stem) | Slim stem and relatively short, dense panicle and high degree of sterile and low seed setting |
| Trisomic VI (dense peduncle) | Short and strong stem, wide, long and dark green leaves with ripples, distorted and curled peduncle, and dense panicle |
| Trisomic VII () | Creep, slim and soft stem with tillers, relatively long leaves, loose panicle with long bristles |
| Trisomic VIII (panicle with narrow top) | Plant stature similar with wild types, not well-arranged branches on the panicle with narrow top, early heading and well seed setting |
| Trisomic IX (normal plant) | Plant stature and panicle morphology similar with wild normal Yugu 1 with relatively good seed setting |

chromosomes instead of the normal two copies, so the karyotype of the series is $2n+1=18+11, +12, +13\ldots+19$. Because there are three copies of specific chromosomes in individual cells, the resulting biochemical reactions cause each trisomic line to exhibit its own particular morphological characteristics, which can be used to identify chromosomes. The characteristics of the nine trisomic foxtail millet lines are provided in Table 5.5. With this system, the *Anai3* foxtail dwarf gene was located on chromosome 3 (Gao et al. 2000) and the starch *Wx* gene was detected on chromosome 4 (Mao et al. 2000a, b). Additionally, genes responsible for male sterility and yellow leaves were located on chromosomes 6 and 7, respectively (Wang et al. 2002). More importantly, the relationships between RFLP molecular linkage groups and individual chromosomes were characterized, which formed the basis of genomic studies involving *Setaria* species (Wang et al. 1998).

Due to the rapid development of foxtail millet as a model system, detailed and accurate deciphering of foxtail millet morphological and agricultural important characters will increase markedly in the near future, which will provide valuable information for plant functional genomics and crop domestication and evolution studies. Recently published identification of heading date-related QTLs (Mauro-Herrera et al. 2013), grain yield-related QTLs (Fang et al. 2016), height, tillering and biomass-related QTL (Mauro-Herrera and Doust 2016), and a large-scale identification of 512 QTLs for 47 botanical and agronomic characters (Jia et al. 2013) have made a strong foundation for further genetic analyses of foxtail millet morphological characters. In addition to morphological characters, identification of biotic- and abiotic-stress related QTLs via association mapping and linkage mapping are currently under way in several labs and relevant results will be published in the near future.

## References

Afzal M, Kawase M, Nakayama H, et al. Variation in electrophoregrams of total seed protein and Wx protein in foxtail millet. In: Janick J, editor. Progress in new crops. Alexandria: ASHS Press; 1996. p. 191–5.

Austin DF. Fox-tail millets (Setaria: Poaceae)—abandoned food in two hemispheres. Econ Bot. 2006;60(2):143–58.

Avdulov NP. Karyo-systematische untersuchungen der familie Gramineen. Bull Appl Bot Suppl. 1931, 43. Leningrad.

Ayyangar GNR, Narayanan TR. The inheritance of characters in *Setaria italica* (Beauv.) the Italian millet, part I. Grain colours. Ind J Agric Sci. 1931;1:586–608.

Ayyangar GNR, Narayanan TR. Inheritance of characters in *Setaria italica* (Beauv.) the Italian millet, part II. Anther colours. Ind J Agric Sci. 1932;2:59–61.

Ayyangar GNR, Narayanan TR. The inheritance of characters in *Setaria italica* (Beauv.) the Italian millet, part III. Bristles. Ind J Agric Sci. 1933;3:207–18.

Ayyangar GNR, Narayanan TR, Rao TN, Sarma PS. Inheritance of characters in *Setaria italica* (Beauv.) the Italian millet, part VII. Plant purple pigmentation. Ind J Agric Sci. 1935;5:176–94.

Bonjean APA. Origins and historical diffusion in China of major native and alien cereals. In: He Z, Bonjean APA, editors. Cereals in China. Mexico, DF: CIMMYT; 2010. p. 1–14.

Chen JG, Zhou X, Zhang YZ. Gibberellin-responding and non-responding dwarf mutants in foxtail millet. Plant Growth Regul. 1998;26(1):19–24.

Cui G. Breeding for disease resistant foxtail millet cultivars. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 447–71.

Cui G, Zheng G, Dong Z. Screening of foxtail millet materials resistant to Aphelenchoides besseyi. Acta Agric Bor Sin. 1989;S1:145–51.

Dekaprelerich L, Kasparian ASA. Contribution to the study of foxtail millet (*Setaria italica* P.B. maxima Alf.) cultivated in Georgia (Western Transcaucasia). Bull Appl Bot Genet Plant Breed. 1928;19:533–72.

d'Ennequin MLT, Panaud O, Toupance B, et al. Assessment of genetic relationships between Setaria italica and its wild relative *S. viridis* using AFLP markers. Theor Appl Genet. 2000;100(7):1061–6.

Diao X. Current status of foxtail millet production in China and future development directions. In: Diao X, editor. Foxtail millet industry and the Chinese foxtail millet research system. Beijing: China Agricultural Scientific and Technology Press; 2011. p. 20–30.

Dineshkumar SP, Shashidhar VR, Ravikumar RL, et al. Identification of true genetic dwarfing sources in foxtail millet (*Setaria italica*, Beauv.). Euphytica. 1992;60:207–12.

Dong L, Ma J, Zheng Z, et al. Identification of millet cultivars resistance to nematode and screening of chemicals for seed dressing. J Hebei Agric Sci. 2010;14(11):54–5.

Dong L, Wang X, Shi A, et al. Identification of rust resistance in main production varieties of millet. J Hebei Agric Sci. 2012;8:1–4.

Dong L, Quan J, Lu P, et al. Identification of main millet cultivars resistance to millet blast. Chin Plant Protec. 2015;35(7):33–7.

Doust AN, Devos KM, Gadberry MD, et al. Genetic control of branching in foxtail millet. Proc Natl Acad Sci U S A. 2004;101(24):9045–50.

Duan C, et al. Effects of selection on economic traits of spring foxtail millet. In: Li D, editor. Breeding technology of foxtail millet cultivars. Xi'an: Press of Tianze; 1990. p. 179–81.

Eda M, Izumitani A, Ichitani K, et al. Geographical variation of foxtail millet, *Setaria italica* (L.) P. Beauv. based on rDNA PCR–RFLP. Genet Resour Crop Evol. 2013;60(1):265–74.

Fang X, Dong K, Wang X, Liu T, He J, Ren R, Zhang L, Liu R, Liu X, Li M, Huang M, Zhang Z, Yang T. A high density genetic map and QTL for agronomic and yield traits in foxtail millet (*Setaria italica* (L.) Beauv.). BMC Genomics. 2016;17:336.

Fukunaga K, Domon E, Kawase M. Ribosomal DNA variation in foxtail millet, *Setaria italica* (L.) P. Beauv., and a survey of variation from Europe and Asia. Theor Appl Genet. 1997; 95(5–6):751–6.

Fukunaga K, Wang Z, Kato K, et al. Geographical variation of nuclear genome RFLPs and genetic differentiation in foxtail millet, *Setaria italica* (L.) P. Beauv. Genet Resour Crop Evol. 2002a;49(1):95–101.

Fukunaga K, Kawase M, Kato K. Structural variation in the Waxy gene and differentiation in foxtail millet [*Setaria italica* (L.) P. Beauv.]: implications for multiple origins of the waxy phenotype. Mol Genet Genomics. 2002b;268:214–22.

Fukunaga K, Ichitani K, Taura S, et al. Ribosomal DNA intergenic spacer sequence in foxtail millet, *Setaria italica* (L.) P. Beauv. and its characterization and application to typing of foxtail millet landraces. Hereditas. 2005;142(2005):38–44.

Fukunaga K, Ichitani K, Kawase M. Phylogenetic analysis of the rDNA intergenic spacer subrepeats and its implication for the domestication history of foxtail millet, *Setaria italica*. Theor Appl Genet. 2006;113(2):261–9.

Fukunaga K, Ichitani K, Kawase M. rDNA polymorphism of foxtail millet (*Setaria italica* ssp. italica) landraces in northern Pakistan and Afghanistan and in its wild ancestor (*S. italica* ssp. viridis). Genet Resour Crop Evol. 2011;58(6):825–30.

Gan Y. Breeding for pest resistant foxtail millet cultivars. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 472–90.

Gao J, Mao L, Wang R. A study on the cytology and morphology of the tetrasomics in foxtail millet. Acta Agron Sin. 2000;26(6):801–5.

Gao JH, Wang RQ, Mao LP, Diao XM. Chromosome locating of dwarf gene in foxtail millet An'ai3. Acta Agron Sin. 2003;29(1):152–4.

Gu S, Liu S, Liu Z. Inheritant effects of grain lipid content in spring-sowing foxtail millet. J Shanxi Agric Sci. 1991;2:5–9 (In Chinese).

Gu S, Liu S, Liu Z. The genetic analysis on protein content in spring foxtail millet hybrids. Acta Agric Bor Sin. 1992;7(3):1–8.

He S, Yang Y, Morrell PL, et al. Nucleotide sequence diversity and linkage disequilibrium of four nuclear loci in foxtail millet (*Setaria italica*). PLoS One. 2015;10(9):e0137088.

Hirano R, Naito K, Fukunaga K, et al. Genetic structure of landraces in foxtail millet (*Setaria italica* (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet. Genome. 2011;54(6):498–506.

Jia G, Huang X, Zhi H, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013;45(8):957–61.

Jia G, Liu X, Schnable JC, et al. Microsatellite variations of Elite Setaria varieties released during last six decades in China. PLoS One. 2015;10(5):e0125688.

Kawase M, Sakamoto S. Geographical distribution and genetic analysis of phenol color reaction in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1982;63(2):117–9.

Kumari R, Dikshit N, Sharma D, et al. Analysis of molecular genetic diversity in a representative collection of foxtail millet [*Setaria italica* (L.) P. Beauv.] from different agro-ecological regions of India. Physiol Mol Biol Plants. 2011;17(4):363–74.

Laboratory of Foxtail Millet, Institute of Crops Sciences, Inner Mongalia. Summary of foxtail millet breeding researches. Agric Sci Exp. 1979;1:27–35.

Laboratory of Millet Crops, Institute of Crops of Hebei Province. A preliminary study on heritability, inheritance correlation and selection index of the main characters of summer millet. Acta Genet Sin. 1975;2(3):249–54.

Li Y. A study on the identification of drought-resistance on millet germplasm. Acta Agric Bor Sin. 1991;3:20–5.

Li Y. Reliability and practicability of repeat drought tolerance identification in foxtail millet at seedling stage. Hebei Agric Sci. 1992;4:9–11.

Li X, Chen R. The karyotype analysis of *Setaria italica* and *S.viridis*. J Wuhan Bot Res. 1985;3(4):409–12.

Li Y, Ling L. Analysis of correlation between quality character and other main agronomic characters of summer-sown foxtail millet. Acta Agric Bor Sin. 1989;S1:44–50.

Li HW, Meng CJ, Liu TN. Problems in the Breeding of Millet (*Setaria Italica* (L.) Beauv.). Agron J. 1935;27(12):963–70.
Li HW, Meng JC, Li CH. Genetic studies with foxtail millet,(*Setaria italica* L.). Beauv. J Am Soc Agron. 1940;32:426–38.
Li HW, Li CH, Pao WK. Cytological and genetical studies of the interspecific cross of the cultivated foxtail millet, *Setaria italica* (L.) Beauv., and the green foxtail millet, *S. viridis* L. J Am Soc Agron. 1945;37(1):32–54.
Li Y, Wu S, Cao Y. Cluster analysis of an international collection of foxtail millet (*Setaria italica* (L.) P. Beauv.). Euphytica. 1995;83(1):79–85.
Li Y, Jia J, Wang Y, et al. Intraspecific and interspecific variation in Setaria revealed by RAPD analysis. Genet Resour Crop Evol. 1998;45(3):279–85.
Li W, Zhi H, Wang Y, et al. Assessment of genetic relationship of foxtail millet with its wild ancestor and close relatives by ISSR markers. J Integr Agric. 2012;11(4):556–66.
Liu W. Estimation of the genetic parameters for main characters of millet and their applications in breeding. J Shanxi Agric Univ. 1984;4(2):172–81.
Liu Z, et al. Correlation analysis of protein and lipid content in F1 hybrids of spring foxtail millet. In: Li D, editor. Breeding technology of foxtail millet cultivars. Xi'an: Press of Tianze; 1990. p. 132–6.
Liu Z, Bai G, Zhang D, et al. Genetic diversity and population structure of elite foxtail millet [*Setaria italica* (L.) P. Beauv.] germplasm in China. Crop Sci. 2011;51(4):1655–63.
Liu Z, Zhang T, Li C, Bai G. Genetic diversity and classification of cytoplasm of Chinese elite foxtail millet [*Setaria italica* (L.) P. Beauv.] germplasm. Crop Sci. 2014;54(2):659–66.
Lu P. Descriptors and data standard for foxtail millet (*Setaria italica* Beauv). Beijing: China Agricultural Press; 2006. p. 1–75.
Lu P. Foxtail millet germplasm and breeding. In: Diao X, editor. Strategic study of durable development of foxtail millet. Beijing: China Agricultural Press; 2016.
Ma J, Dong L, Gao L, Gan Y, Dong Z. Identification of millet variety resistance to eyespot and innovation of resistance germplasm. Zaliangzuowu. 2005;25(6):389–91.
Mao L, Gao J, Wang R, Sun L. Chromosome location of glutinous trait gene of foxtail millet endosperm. Acta Agric Bor Sin. 2000a;15(4):10–4.
Mao L, Gao J, Wang R. Chromosome location of glutinous trait gene of foxtail millet endosperm. Acta Agric Bor Sin. 2000b;15(4):10–3.
Mauro-Herrera M, Doust AN. Development and genetic control of plant architecture and biomass in the panicoid grass, *Setaria*. PLoS One. 2016;11(3):e0151346.
Mauro-Herrera M, Wang X, Barbier H, et al. Genetic control and comparative genomic analysis of flowering time in Setaria (Poaceae). G3 Genes Genom Genet. 2013;3(2):283–95.
Ochiai Y. Variation in tillering and geographical distribution of foxtail millet (*Setaria italica* P. Beauv.). Breeding Sci. 1996;46(2):143–8.
Qian J, Jia G, Zhi H, Li W, Wang Y, et al. Sensitivity to gibberellin of dwarf foxtail millet (*Setaria italica* L.) varieties. Crop Sci. 2012;52:1068–75.
Rao KEP, De Wet JMJ, Brink DE, et al. Infraspecific variation and systematics of cultivated *Setaria italica*, foxtail millet (Poaceae). Econ Bot. 1987;41(1):108–16.
Schontz D, Rether B. Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv.-RFLP using a heterologous rDNA probe. Plant Breed. 1998;117(3):231–4.
Schontz D, Rether B. Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv.: identification and classification of lines with RAPD markers. Plant Breed. 1999;118(2):190–2.
Sun P, Qiu Y, Guo X, et al. Cytological study of foxtail millet, I. Karyotye and Giemsa C banding analysis. J Shanxi Agric Sci. 1983;6:13–5.
Sun P, Zhou X, Hou B, et al. A cytogenetical study of *Setaria italica* VI. Preliminary study on origin. Hereditas. 1994;16(3):24–7.
Takahashi N. Studies on the flower of Italian millet and a method of its artificial hybridization (in Japanese). Proc Crop Sci Soc Jpn. 1942;13:337–40.
Upadhyaya HD, Pundir RPS, Gowda CLL, et al. Establishing a core collection of foxtail millet to enhance the utilization of germplasm of an underutilized crop. Plant Genet Resour. 2009;7(02):177–84.

Upadhyaya HD, Ravishankar CR, Narasimhudu Y, et al. Identification of trait-specific germplasm and developing a mini core collection for efficient use of foxtail millet genetic resources in crop improvement. Field Crop Res. 2011;124(3):459–67.

Wang R, Gao J, Wang Z, et al. Establishment of trisomic series of millet (*Setaria italica* [L.] Beauv.). Acta Bot Sin. 1994;36(9):690–5.

Wang Z, Devos K, Liu C, et al. Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P Beauv. Theor Appl Genet. 1998;96:31–6.

Wang R, Gao J, Liang X. Identification of primary trisomics and other aneuploids in foxtail millet. Plant Breed. 1999;118:59–62.

Wang R, Gao J, Mao L, et al. Chromosome location of the male-sterility and yellow seedling gene in line 1066 of foxtail millet. Acta Bot Sin. 2002;44(10):1209–12.

Wang Y, Zhi H, Li W, et al. A novel genome of C and the first autotetraploid species in the *Setaria* genus identified by genomic in situ hybridization (GISH). Genet Resour Crop Evol. 2009;56(6):843–50.

Wang C, Chen J, Zhi H, et al. Population genetics of foxtail millet and its wild ancestor. BMC Genet. 2010;11(1):90.

Wang C, Jia G, Zhi H, et al. Genetic diversity and population structure of Chinese foxtail millet [*Setaria italica* (L.) Beauv.] landraces. G3 Genes Genom Genet. 2012;2(7):769–77.

Wang H, Jia G, Zhi H, Wen Q, et al. Phenotypic diversity evaluations of foxtail millet core collections. Acta Agronomica Sinica. 2016. http://www.cnki.net/kcms/detail/11.1809.S.20151008.1403.032.html

Xue C, Zhi H, Fang X, et al. Characterization and fine mapping of *SiDw2* in foxtail millet. Crop Sci. 2016;56(1):95–103.

Yao Z, Liang T. The development of a dominant dwarf variety of foxtail millet. In: Li D, editor. Advances in foxtail millet breeding. Xi'an: Tianze Press; 1990. p. 82–6.

Zhang L. Crossing based pedigree breeding of foxtail millet. Shanghai: Shaihai Science and Technology Press; 1961.

Zhang L. Pedigree selection of foxtail millet. Shanghai: Press of Science and Technology of Shanghai; 1972.

Zhang W, Zhi H, Liu B, et al. Indexes screening for drought resistance test of foxtail millet. J Plant Genet Resour. 2010;11(5):560–5.

Zhang W, Zhi H, Liu B, et al. Effects of drought stress on millet photosynthetic characteristic in booting stage. J Hebei Agric Sci. 2011a;15(6):7–11.

Zhang W, Zhi H, Liu B, et al. Relationship between drought tolerance ability and some physiological characteristics of foxtail millet at booting stage. Acta Agric Bor Sin. 2011b;26(3):128–33.

Zhang W, Zhi H, Liu B, et al. Screening of indexes for drought tolerance test at booting stage in foxtail millet. J Plant Genet Resour. 2012;13(5):765–72.

Zhao H, Jin L. Variations and inheritance of quantitative traits in several type of foxtail millet cultivars. In: Academy of Agricultural Sciences in Heilongjiang of China, editor. Alum of foxtail millet studies in Northeast China. Heilongjiang: Academy of Agricultural Sciences; 1985. p. 89–92.

Zhao M, Zhi H, Doust AN, et al. Novel genomes and genome constitutions identified by GISH and 5S rDNA and Knotted 1 genomic sequences in the genus *Setaria*. BMC Genomics. 2013;14:244.

Zhao M, Zhi H, Zhang X, Jia G, Ma J, Li Y, Liu X, Diao X. Identification and characterization of Setaria DELLA protein reveal its gibberellin (GA) independent interaction with SiGID1. New Phytologist 2016; submitted in review.

Zhi H, Diao X, Lv F, et al. Methodology analysis on screening of salt tolerant genotypes from foxtail millet and other Setaria species. J Hebei Agric Sci. 2004;4:15–8.

# Chapter 6
# Foxtail Millet Breeding in China

**Xianmin Diao and Guanqing Jia**

**Abstract** Although there is a long history of foxtail millet cultivation in China, modern foxtail millet breeding was only initiated in China in the 1950s and 1960s, with significant progress being made since the 1980s. Most of the research on foxtail millet breeding has been conducted in China, where it is an important regional cereal. The main research activities from the 1950s to 1970s were comparisons among landraces and individual selection, followed by cross-based pedigree selection in the 1970s. These comparisons and cross-based pedigree selections contributed greatly to foxtail millet improvement in China, including the development of the super cultivars 'Yugu 1' and 'Zhaogu 1' in the 1980s. Radiation and chemical-induced mutations have also been used in foxtail millet breeding to create novel types, such as dwarf lines. Although different types of male sterile lines have been developed over the past 50 years in China, only partial genetic male sterile lines (PAGMS) have been used successfully in hybrid seed production, allowing the use of heterosis to become a reality in recent years. The foxtail millet eco-regions, breeding phases, breeding methodology, and main cultivars grown at different times since the 1950s in China are reviewed in this chapter. With the rapid advances in foxtail millet genomic sciences, mining and elucidation of quantitative trait loci related to important traits will accelerate foxtail millet breeding in the near future.

**Keywords** Foxtail millet • Pedigree selection • Cultivar development • Heterosis • Setaria

X. Diao, Ph.D. (✉) • G. Jia
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences,
No.12, Zhongguancun South St., Haidian District, Beijing 100081,
People's Republic of China
e-mail: diaoxianmin@caas.cn



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_6

## 6.1 Foxtail Millet Breeding in China

### *6.1.1 Eco-Regions and Ecotypes of Foxtail Millet in China*

Foxtail millet is an ancient crop in China, where it has probably been cultivated for 10,000 years. Under long-term selective pressures in different geographical, climatic, and anthropogenic regions, foxtail millet has evolved into diverse ecotypes for field production in different regions. Each ecotype includes many varieties with common genetic characteristics that allow them to adapt to certain ecological conditions, including differences in photoperiod. The ecological classification of foxtail millet is pivotal to understanding the characteristics of different groups of varieties and to establish and exchange germplasm collections. Studies on the ecological responses of foxtail millet varieties have been concentrated in the main foxtail millet production areas of northern China. Several different methods for classifying eco-regions have been reported, including classification according to the natural climatic conditions of the different foxtail millet-producing areas, and the heading dates of various groups of foxtail millet varieties (Wang and Guo 1997; Diao et al. 2011). The 'four ecological regions' theory was developed based on these classifications, and this has been largely adopted by Chinese foxtail millet breeders. The four ecological regions of foxtail millet are the Northeast Plain, North China Plain, Inner Mongolia Plateau, and the Northwest Plateau.

#### 6.1.1.1 The Northeast Plain Eco-Region

This ecological zone includes Heilongjiang, Jilin, and the northern part of Liaoning Province in northeast China, which are the northernmost regions of China (latitude ≥40°). In this eco-region, the altitude ranges from about 50 to 200 m, the frost-free period is short, and it is mainly the spring-sowing region for foxtail millet cultivation, except for a few zones located in the south of Liaoning. The annual rainfall in this region is about 400–700 mm, and rainfall is concentrated during summer during the growing season of foxtail millet. The accessions in this region have medium–tall plant height (average height, 141.8 cm), medium panicle length (22.6 cm), an average 1000-grain weight of 2.77 g, and an average growing period of about 110–130 days. Recent simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) data have shown that accessions from Heilongjiang Province are greatly diverged from those in Jilin and Liaoning provinces. The varieties from Heilongjiang were shown to be more light- and temperature-sensitive, and consequently, were classified as the early spring-sowing region type. The varieties from Jilin and Liaoning provinces were shown to be genetically much closer to the summer-sowing types cultivated on the North China Plain (Wang et al. 2012; Jia et al. 2013).

#### 6.1.1.2 The North China Plain Eco-Region

This eco-region covers the area from the south of the Great Wall to the north of the Huaihe River, to the west of the Taihang Mountains and most regions in Hengshan distributed in Henan, Hebei, Shandong, Anhui, Jiangsu, Beijing, Tianjin, and other provinces or cities (latitude ≤40°). In this eco-region, the average altitude is about 100 m, and the frost-free period is about 6–8 months. The annual rainfall is about 500–700 mm, with rainfall concentrated during the foxtail millet growing period in summer. This area is appropriate for both spring- and summer-sowing types of foxtail millet. The average growing period of spring-sowing varieties is about 110–120 days and that of summer-sowing varieties is about 80–90 days. The accessions in this region have a medium plant height (average, 127.6 cm), medium to small panicles, and relatively low average 1000-grain weight (2.74 g) compared with those of other ecotypes. Foxtail millet was a spring-sowing crop in this region before the 1960s, but became a summer-sowing crop after the harvesting of wheat from the 1970s onward.

#### 6.1.1.3 Inner Mongolia Plateau Eco-Region

This eco-region includes Hohhot and Chifeng in Inner Mongolia, Fuxin and Chaoyang in Liaoning, Chengde and Zhangjiakou in Hebei, Yanbei in Shanxi, and the Yulin region in Shannxi Province. This ecological zone is located at high latitudes, with an average altitude of 1000 m. The annual rainfall in this region is about 400 mm, and the frost-free period is about 5 months. This area is appropriate for spring-sowing varieties, with an average growing period of 110–120 days. The foxtail millet varieties from this region have an average plant height of 135 cm, large panicles (22.9 cm long), and large grains (average 1000-grain weight, 3.52 g).

#### 6.1.1.4 Northwest Plateau Eco-Region

This eco-region includes Gansu, Shannxi, Xinjiang, and the southern part of Shanxi Province. In this eco-region, the average altitude is about 1000–1500 m, the annual precipitation is about 400–500 mm, and the frost-free period is about 150–180 days. This area is mainly appropriate for spring-sowing varieties with a growing period of about 110–130 days. Some locations in this region are also appropriate for summer-sowing types with a growing period of 85–95 days. The foxtail millet varieties in this region have a medium plant height (average, 135.5 cm), medium-long panicles (average, 29.3 cm), and medium-large grains (average 1000-grain weight, 3.44 g). Recent morphological and molecular studies have identified that varieties from Gansu Province are much more light- and temperature-sensitive than those from other provinces in this region and are genetically similar to varieties from Heilongjiang Province; consequently, they should be classified as the early spring-sowing region type (Wang et al. 2012; Jia et al. 2013).

### *6.1.2 A Brief Introduction to Modern Foxtail Millet Breeding in China*

Modern foxtail millet breeding started in the 1950s and 1960s, and significant progress in foxtail millet cultivar improvement has been made since 1980s. Several books or chapters in Chinese have summarized the progress of foxtail millet breeding, including "Foxtail Millet Breeding" (1997) edited by Li Yinmei, "Principles of Foxtail Millet Genetics and Breeding" by Li (1997), "Genetic Improvement of Foxtail Millet" by Diao et al. (2011), and "Foxtail Millet Production and Research System in China" by Diao (2011). The development of foxtail millet breeding programs in China can be divided into three phases: the direct reselection phase, the pedigree cross-breeding phase, and the multiple-methods-integrated breeding phase.

#### 6.1.2.1 Direct Reselection Phase

Modern improvement of foxtail millet varieties began in the 1920s at several institutes such as the former Jinling University in Nanjing, the Yanjing Crop Improvement Station in Beijing, and the North China Agricultural Research Institute in Beijing. Researchers at these institutes started to evaluate and compare genotypes, leading to the release of some outstanding cultivars such as 'Yanjing 811,' 'Kaifeng 48,' and 'Huanong 4.' During the 1950s–1960s, direct selection from populations with natural variation was the main method to improve foxtail millet. The collected landraces were also extensively compared during the 1950s, and superior varieties were selected. The average grain yield of foxtail millet was about 915 kg/ha during the 1920s to the 1950s in China.

#### 6.1.2.2 Pedigree Cross-Breeding Phase

The period from the 1960s to the 1970s can be described as the pyramiding cross-breeding phase. The Xinxiang Agricultural Research Institute in Henan Province initiated cross-breeding and released the cultivar 'Xinnongdong 2' in 1959. In this method, hybrids are produced by crossing different varieties and super individuals or lines are screened from their offspring. After this period, cross-breeding became popular throughout the country and it is still the main method of foxtail millet breeding today. Since the 1980s, cross-breeding has continued to make substantial contributions to foxtail millet improvement in China. Many leading cultivars developed during this period include 'Yugu 1' and 'Zhaogu 1,' which were developed in the 1980s by the Anyang Agricultural Research Institute in Henan Province and the Chifeng Agricultural Research Institute in Inner Mongolia, respectively. Thereafter, these two varieties became landmark cultivars contributing to national foxtail millet production in China after the 1980s. Many newly developed cultivars are derivatives of these two landmark cultivars. The grain yield of foxtail millet in China reached 1485–1852.5 kg/ha during this period.

#### 6.1.2.3 Integrated Breeding Phase

Since the 1960s, mutations caused by radiation and chemical treatments have been widely used for foxtail millet breeding. In the 1980s, somaclonal variations arising in tissue culture were also exploited for foxtail millet breeding. In the integrated breeding phase, artificial mutations were integrated with hybridization-based pedigree selection to improve foxtail millet. Mutation breeding began in 1963 at the Zhangjiakou Baxia Agricultural Research Institute in Hebei province. The cultivar 'Zhangnong 1' was developed using $^{60}$Coγ-ray mutagenesis. Beginning in the 1970s, mutation breeding was widely used for foxtail millet breeding in China, and this method was responsible for approximately 30 % of all cultivars developed over the following 20 years (Yi 1997). Somatic variations induced by tissue culture were also exploited to improve foxtail millet (Diao et al. 2002), resulting in several cultivars, including 'Ai 88,' that were released in China. Sterility of the $F_1$ progeny resulting from interspecies hybridization has also been relieved through the use of tissue culture (Luo et al. 1993; Zhou et al. 1988). During this period, the average grain yield of foxtail millet in China increased to 2250–3000 kg/ha[2].

#### 6.1.2.4 Release of the Super Cultivars 'Yugu 1' and 'Zhaogu 1' in China

In the 1980s, the grain yields of summer-sown foxtail millet varieties in northern China were extremely low because of lodging and leaf diseases. The cultivar 'Yugu 1' was developed by the Anyang Institute of Agricultural Science in Henan province in the early 1980s by pedigree selection from the offspring of a cross between the land races Riben 60 Days and Tulong. In 1983, 'Yugu 1' was authorized and released by the Crop Assessment Committee of Henan Province. This variety exhibited excellent agronomic traits, including lodging resistance, a high production rate of effective tillers (over 95 %), and high grain yield. 'Yugu 1' represented a significant breakthrough in combining grain yield and taste quality. During a 2-year field trial in 1985 and 1986 in Xinjiang Hami City, 'Yugu 1' was identified as a first-order drought-tolerant variety. This variety also showed resistance to valley blast, brown stripe disease, brown spot, and smut disease. 'Yugu 1' also showed wide adaptability to environmental conditions, and it has been cultivated in 24 provinces of China since its release. Ever since the 1980s, 'Yugu 1' has been used frequently as a parent in hybrid-based pedigree selection breeding. Currently, nearly 80 % of the cultivars released in the summer-sowing regions of China are direct or indirect derivatives of 'Yugu 1.' Today, 'Yugu 1' is not only widely cultivated in China, but it is also used extensively in breeding and basic research. For example, it was used to generate the foxtail millet trisomic series (Wang et al. 1994) and the foxtail millet reference genome (Bennetzen et al. 2012).

Around the same time in the 1980s, 'Zhaogu 1' was released by the Chifeng Institute of Agricultural Sciences in eastern Inner Mongolia. This variety showed outstanding lodging resistance, thus solving the lodging problem for spring-sowing

foxtail millet varieties. Compared with conventional foxtail millet varieties, 'Zhaogu 1' was shown to be more resistant to white hair diseases. Consequently, this variety led to a new generation of high- and stable-yielding and adaptable varieties in spring-sowing foxtail millet breeding.

#### 6.1.2.5 Breeding for Taste and Nutritional Quality

As well as higher grain yields, superior taste and nutritional properties are breeding targets in Chinese foxtail millet breeding programs. In the long history of foxtail millet cultivation, many varieties have been selected for their superior taste, including the four well-known Chinese varieties 'Qinzhouhuang,' 'Jinmi,' 'Longshanmi,' and 'Taohuami.' Although several researchers have tried to establish breeding methods to improve taste and nutritional quality over the past 30 years (Gao and Wang 1997), little is known about the heredity of these characters, which hampers effective breeding.

Starting from the national "Seventh-Five-Year Plan" in the 1980s, taste and nutritional quality were established as breeding targets and several high-quality and special-use cultivars were released. For instance, 'Jingu 21' with superior commercial quality was released by the Institute of Economical Crops, Agricultural Academy of Shanxi Province, and 'Xiaoxiangmi,' 'Jiyou 2,' and 'Jite 4' were released by the Institute of Millet Crops, Hebei Academy of Agricultural and Forestry Sciences. Since 1988, the Millet Section of the Crop Science Society of China continued a national assessment of foxtail millet grain quality, focusing on commercial use and superior taste, and more than 80 varieties were designated as good-quality foxtail millet cultivars. In the past 60 years, approximately 770 varieties have been registered or released in China, but only a small subset are still cultivated. Table 6.1 lists the leading cultivars grown in China from the 1950s to the present day.

### *6.1.3 Foxtail Millet Breeding in Countries Other than China*

The main foxtail millet breeding programs in countries other than China are in India and France. In India, breeding of foxtail millet cultivars began in the 1930s, and the cultivars released from the 1970s to the 2000s included the 'Co' series (Coinbatore), the 'HK' series (Bangalore), the 'K' series (Bangalore), 'Arjuna,' 'SIA326,' 'Chitra' (Andhra Pradesh), the 'SIA' series (Bihar), 'K221' (Karnataka), 'Rs118' (Karnataka), and the 'SIC' series (Maharashtra). Most of the foxtail millet breeding programs in Korea and Japan were conducted in the early 1900s, and almost no breeding has been conducted in those countries since the 1970s. In France, foxtail millet was improved for bird-feed production, and herbicide-resistant green foxtail was used to breed the herbicide-resistant foxtail millet varieties that were introduced into China in the late 1990s (Wang et al. 2001) (also see Chap. 15). Work related to foxtail millet breeding has also been conducted in America, Russia, and some other countries.

**Table 6.1** Leading cultivars of foxtail millet from the 1950s to the present

| Cultivar | Cross | Year of release | Areas of adaptation |
|---|---|---|---|
| Xinnong 724 | Mihuanggu reselection | 1955 | Henan, Hebei, Shandong |
| Moligu | Jiansuijinmiaohuang reselection | 1958 | Hebei, Beijing, Shanxi |
| Hualian 1 | Huaidehualiangu reselection | 1959 | Jilin |
| Angu 18 | Daqingmiao reselection | 1965 | Heilongjiang |
| Lugu 2 | 60 days Huancang reselection | 1971 | Shandong |
| Jingu 1 | Baimujizui reselection | 1973 | Shanxi |
| Zhaogu 1 | Shuangguayin reselection | 1977 | Inner Mongolia, Liaoning |
| Baisha 971 | Baishagu reselection | 1978 | Jilin |
| Jigu 6 | Japan 60 days × Xinnong 724 | 1982 | Hebei, Henan, Shandong |
| Yugu 1 | Japan 60 days × Tulong | 1983 | Henan, Hebei, Shandong |
| Longgu 25 | Harbin 5 × Longgu 23 | 1986 | Heilongjiang |
| Yugu 2 | An 30 × Xiaoliugen | 1989 | Henan, Hebei, Shandong |
| Jingu 21 | Jinfen 52 mutation | 1991 | Shanxi, Shanxi, Inner Mongolia |
| Lugu 10 | Yugu 1 × 5019-5 | 1995 | Shandong, Hebei |
| Jigu 14 | Lusuigu mutation | 1996 | Hebei, Henan, Shandong |
| Gufeng 2 | 95307 × 8337 | 2002 | Hebei, Henan, Shandong |
| Jigu19 | Ai88 × Qingfenggu | 2004 | Hebei, Henan, Shandong |
| Yugu 18 | Yugu1 × Bao282 | 2012 | Henan, Hebei, Shandong, Liaoning |
| | | | Xinjiang, Beijing |

From Cheng and Dong (2010), with some modifications

## *6.1.4 Conventional Foxtail Millet Breeding Technologies*

### 6.1.4.1 System Selection

Natural mutations occur frequently in cultivated foxtail millet, and variations can also arise from hybrids derived from outcrossing. Therefore, it is possible to develop new cultivars from direct selection from the farm field. At the early breeding stage in the 1950s, direct selection was popular in China, and even today it is a simple and effective method for breeders.

There are two options for direct selection: single plant selection and pooled selection. For single plant selection, healthy panicles from well-developed robust plants with characters matching breeding targets are selected, numbered, and preserved. Each panicle becomes a line in the next generation, and different lines are compared against the original varieties. Then, conventional pedigree selection over the following years results in lines with superior characters. 'Gonggu 6,' 'Changnong 1,' 'Hengyan 130,' 'Lugu2,' and other cultivars were developed using this method. These cultivars showed 5–30 % yield increases compared with those of the original varieties. The second method, pooled selection, is when varieties cultivated for a long time are classified into types based on their variations. For pooled selection, plants with similar

characters are harvested together as a pool. Then, in the following year, comparisons are made among the pools, the original variety, and the check variety in field trials, and superior varieties are selected. The foxtail millet cultivars 'Huanong 4,' 'Baisha 971,' 'Changwei 69,' 'Lugu 4,' and others were developed by the pooled method. Sometimes the pooled method can be used in combination with the single plant selection method, depending on the degree of variation in the basic varieties.

#### 6.1.4.2 Cross-Based Pedigree Selection

Cross-based pedigree selection can combine superior characters from different varieties, and it is the most popular method for foxtail millet improvement. The choice of the parents and the selection of the best individuals and lines from the offspring are the key steps for successful pedigree selection.

##### Hybrid Parent Selection and Arrangement

Only those varieties with many good traits and few unfavorable ones can be used as parents. The two parents in a given cross should have complementary characters, especially the main agriculturally important characters. The female parent, which is usually a local cultivated variety, cannot have serious defects. Varieties with weak photoperiod-sensitivity are usually more adaptable and are good options as parents. The parents can have common advantages, but they should not have common weaknesses. The breeding targets can be very different for different regions and purposes, so parent selection and arrangement can differ widely depending on the local breeding targets.

##### Method of Hybridization

Because the spikelets of foxtail millet are very small and the flowering time differs among spikelets on the same panicle, it is difficult to hybridize different foxtail millet varieties. Chinese breeders have created various methods of hybridization, whose labor inputs and effectiveness differ widely (Cheng and Liu 1997). The first is artificial emasculation of the foxtail millet spikelet with scissors and tweezers before flowering, followed by pollination with male parent pollen. This method is labor intensive and time consuming, and it is rarely used today. The second, frequently used, method, is hot water emasculation. In this method, the female panicle is soaked in water at 47 °C for 8–10 min and then bagged with male panicles for pollination. In the third method, the female panicle or plants are sealed in plastic bags in the field for 2–3 days, depending on the environmental conditions. The high temperature inside the bag kills the pollen, allowing for subsequent pollination with the desired pollen. Various chemicals to kill pollen were also tested for foxtail millet emasculation. Although some were effective, they are rarely used in breeding because of their limited availability.

All of these foxtail millet emasculation methods have a common problem; that is, the kill rate of the pollen in the female panicle is not 100 %, and so genuine hybrids must be identified in the next generation. During breeding, the male parent is often a variety with a dominant character, for example, purple vs. green seedling sheath, green vs. yellow leaf, or orange vs. white anther. In the $F_1$ generation, individuals with the dominant characters are usually genuine hybrids that can be used for subsequent breeding activities. Some quantitative characters can also be used for hybrid identification. For example, early heading and tillering are usually dominant traits, and the $F_1$ hybrids could have heading dates similar to that of the early heading parent or between those of the two parents.

### Pedigree Selection of Hybrid Offspring

Segregation in the second-generation ($F_2$) is complex, especially for those hybrids whose parents are very different phenotypically and genetically, or whose parents are from different ecotype origins. The population size of the $F_2$ generation depends on the degree of difference between the parents, and a larger population size is required for hybrids of vastly different parents and/or those with characters that have a low frequency of recombination, to increase the chances of selecting targeted individuals. Breeding experience has verified that the selection of heading date, maturity date, disease and insect resistance, plant height, grain color, kernel color, and stem strength (lodging resistance) in the $F_2$ generation is effective, because those characters are easy to select for in early generations. Relatively larger numbers of individuals with different combinations of characters should be selected from hybrids with superior parents.

In the $F_3$ generation, most agronomically important characters are still segregating, and individuals with new combinations of different traits are also emerging. In some lines, the degree of segregation in the $F_3$ generation seems to be smaller than that in the $F_2$ generation. For this generation, the selection focus should be on characters with complex genetic backgrounds such as yield and nutritional characters. Individual plant selection should be carried out to obtain lines with superior characters that are still segregating. Those lines that do not match the breeding target should be rejected. The $F_2$ and $F_3$ generations are the key generations for pedigree selection, and it is important to establish appropriate selection criteria for individual plants and to keep a sufficient but not an excessive number of lines. It is important to select individual plants according to the established objectives.

From the fourth to the seventh generations, more and more lines become genetically stable with the increasing number of generations, and the degree of segregation becomes smaller. The selection focus is to identify superior lines that match the breeding targets and to select lines that are still segregating. Comprehensive comparisons among lines should focus on the breeding targets as criteria, and only a small number of lines matching the breeding targets with superior characters and stable inheritance should be selected for the final test. Most other lines will be abandoned at this point. If some important traits are still segregating in a superior line, individual plant selection is still necessary.

The final test of the selected lines is to compare grain yield and quality among different locations and different years. Only those lines with superior characteristics and wide adaptability can be registered as new cultivars and released to farmers.

Breeding experience has verified that the grain yield of foxtail millet is positively correlated with grain weight of the main panicle and flag leaf area, and the straw yield is positively correlated with plant height and growth duration. The taste and nutritional qualities are usually tested at higher generations. Protein and fat contents are positively related to seed size and negatively related to panicle length. Those corelationships can be used to guide the selection of individuals and lines.

Foxtail millet is sensitive to photoperiod, temperature, and other environmental factors, and so it is important to grow hybrid offspring in different environments to select for wide adaptability. Lines with few morphological changes in different environments usually have wider adaptability. For successful breeding, the nursery field should be a little more fertile than the targeted extension field.

#### 6.1.4.3 Breeding Using Radiation and Chemical-Induced Mutations

The roles of radiation and chemical-induced mutations in the breeding of foxtail millet have been summarized by Yi (1997). Because dry seeds are easy to store, transport, and treat, they are usually the materials used for radiation breeding. However, wet seeds, pollen, and even entire plants can be subjected to radiation to induce mutations. There are external and internal radiation treatments for seeds. The external radiation treatment is usually $^{60}$Co γ rays irradiated onto dry seeds with a 30,000–40,000 roentgen dosage, as dosages higher than this are lethal to seeds. If the seed moisture content is high, the treatment dosage is lower. For internal irradiation, seeds of foxtail millet are generally soaked in solutions of chemicals containing $^{32}$P or $^{35}$S. When these elements are absorbed, they emit radiation and cause DNA mutations in the treated plants. Other chemicals have also been used to induce DNA mutations in foxtail millet, including ethyl methanesulfonate (EMS) (Li et al. 2015) and other chemicals (Yi 1997).

The treated seeds and the plants obtained from them are the $M_1$ generation, and usually there is no selection in this generation. However, if there is clear segregation, some individual plants can be selected. The $M_1$ seeds and plants must be cultivated carefully in optimum conditions. After harvesting, 15–20 seeds are collected from each of the $M_1$ individual plants, so as to form a mixed pool designated as $M_2$. The number of seeds collected from each $M_1$ plant varies depending on the number of $M_1$ plants and the size of the $M_2$ pool that have been specified in the experimental design. Because of genetic segregation in the $M_2$ generation, this generation must be observed carefully throughout the whole growing period, in order to select individuals with characters matching the breeding objective. A large number of plants should be selected for those pools with fine characters and wide segregation, and pools with no segregation and/or with bad agronomic phenotypes should be discarded. The selected $M_2$ individuals are grown as lines in the $M_3$ generation. In this generation, the selection should focus on line identification, and individual plants should be selected from the best-performing lines. Because most lines become genetically stable in the $M_4$

generation, the main task for this generation is to comprehensively assess the lines. Lines with good characters matching the breeding targets are selected based on yield and other characters for the final regional adaptation trials.

Most radiation-induced mutations are recessive mutations and can only be selected in the homozygous form after the $M_3$ generation. However, many researchers have found dominant mutations in the $M_1$ generation, and these can be selected in the early generations. So far, more than 40 foxtail millet cultivars have been developed by radiation breeding (Yi 1997) and EMS-induced mutation has played an important role in functional genomics research (Li et al. 2015, 2016).

#### 6.1.4.4 Polyploid Breeding

The ploidy levels of *Setaria* species vary, but domesticated foxtail millet is a true diploid. Because tetraploid species usually have larger seeds, scientists have been very interested in the breeding of tetraploid foxtail millet. In the 1960s and 1970s, polyploid breeding programs developed some tetraploid foxtail millet lines such as 'Wulijin,' 'Jiaqihuang,' and 'Chaoyanggu.' Also, a few foxtail millet landraces with large seeds were identified as natural tetraploids.

Compared with diploid lines, tetraploid foxtail millet lines have shorter, thicker, wider leaf blades, and the blade surface is wrinkled. They also have larger spikelets, pollen grains, and seeds, later heading dates, and longer growth duration. The tetraploid plants are shorter than diploid ones and have smaller panicles. The seed setting rate of the tetraploid foxtail millet is low, leading to low grain yield (Ahanchede et al. 2004). The low grain yield is the main reason why no tetraploid foxtail millet cultivars have been released so far.

However, there are other purposes for tetraploid foxtail millet besides grain production. Tetraploid foxtail millet can be used for distant hybridizations with other tetraploid *Setaria* species such as *Setaria faberii* (Chen et al. 1997; Wu and Bai 2000) and *Setaria yunnanensis* (Zhou et al. 1988). Tetraploid foxtail millet has also been used to construct the foxtail millet trisomic system and in other genetic research (Wang et al. 1994).

There are several different ways to develop tetraploid foxtail millet. The most popular method is to treat seeds with an aqueous solution of colchicine (0.02–0.05 % w/v) for 3–5 days. Other methods include variable-temperature processing of panicles and treatment of young seedlings or tissue-cultured calli.

#### 6.1.4.5 Distant Hybridization

Distant hybridization between different species and/or between different genera can transfer genes from wild to domesticated cultivars. Because foxtail millet is a relatively regional crop, there are only a few research groups that specifically focus on it. Therefore, few distant hybridization breeding and related studies have been conducted for foxtail millet. The earliest report of distant hybridization in foxtail millet was in 1942 by Li et al., who described the hybridization of foxtail millet with green

foxtail and *S. faberii*. Based on this study, they identified that green foxtail is the ancestor of foxtail millet. In the 1960s to 1990s, there were many attempts to hybridize foxtail millet with green foxtail, *S. faberii*, *Setaria glauca*, *S. yunnanensis*, *Setaria verticillata,* and other species, with the goal to develop cytoplasmic male sterile (CMS) lines. Although many distant hybridization attempts were made, real hybrids were only obtained from crosses between foxtail millet and green foxtail, *S. faberii*, *S. verticillata,* and *S. yunnanensis* (Zhu et al. 1987; Chen et al. 1997; Zhou et al. 1988; Wu and Bai 2000; Zhi et al. 2007). Since 2000, distant hybridizations were successfully used to transfer herbicide resistance genes from wild green foxtail into domesticated foxtail millet (Chap. 15). The wild germplasm of *Setaria* is certainly a rich gene pool for the improvement of foxtail millet. Therefore, it is reasonable to expect that distant hybridizations will be a powerful tool to improve foxtail millet in the future.

Because of the close relationship between foxtail millet and green foxtail, crosses between these two types cannot be regarded as distant hybridization. They produce a fertile hybrid, which makes it possible to construct recombinant inbred lines for domestication-related studies and other developmental research, such as the characterization of the foxtail millet root system (Zhang et al. 2014).

## 6.2 Heterosis Utilization in Foxtail Millet

### *6.2.1 Brief History of Research on Foxtail Millet Heterosis Utilization in China*

Heterosis, where hybrids show better growth and grain yield than their parents, is a common phenomenon in crops. Because foxtail millet is a self-fertilizing species with a small spikelet, hybrid seeds can only be produced with the development and use of male sterile lines. The successful use of cytoplasmic male sterile lines in the breeding of rice and other crops prompted foxtail millet breeders to try to develop CMS lines during the 1960s, when the foxtail millet heterosis-utilization breeding program started. The first male sterile foxtail millet was reported by Takahashi (1942) from Japan. In China, the first male sterile line 'Yanxing' was reported in 1967 by Zhu Guanqin and was developed at the Yan'an Agricultural Research Institute of Shannxi Province (Zhu et al. 1991). Since then, many male sterile materials have been identified in farm fields, but none has been successfully developed into a CMS line.

Researchers attempted to develop CMS lines via distant hybridizations of foxtail millet with other *Setaria* species in the 1970s and 1980s. Only one CMS line was produced, and it was never used in hybrid seed production (Zhu et al. 1991). In the early 1970s, the first partial genetic male sterile (PAGMS) line, 'Suanxi 28,' was developed by Cui Wensheng at the Baxia Institute of Agricultural Science, Hebei province. This led to the successful utilization of heterosis in foxtail millet with a two-line system (Cui et al. 1979). In this system, the male sterile line retains

approximately 3 % male fertile spikelets, which can reproduce the male sterile line, and hybrid seeds are produced via hybridization of the PAGMS line with restoration lines. A foxtail millet hybrid cultivar produced using this system, 'Suanxi 28' × 'Zhangnong 10,' was released in the late 1970s. Even today, the two-line system is the only one used for foxtail millet hybrid seed production. In the 1980s, a genetic dominant male sterile line was developed via hybridization between the landraces Aodaliyagu and Tulufan, and a hybrid seed production system was suggested with this dominant male sterile line (Hu et al. 1986, 1993).

In the 1990s, many foxtail millet breeders in China focused on developing photothermo sensitive genetic male sterile (PTGMS) lines, and although several PTGMS lines were developed during this period (Zhao et al. 1994, 1996; Wang et al. 2003), none of them has been used successfully for hybrid seed production. In the first decade of this century, the development of herbicide-resistant restoration lines made the PAGMS two-line system a great success in foxtail millet heterosis utilization, and several new hybrid cultivars were released, including 'Zhangzagu 5,' 'Zhangzagu 8,' and 'Changza 2.' Compared with conventional varieties, hybrid foxtail millet varieties show greatly increased grain yield in suitable cultivation regions, with reported grain yields as high as 7500 kg/ha.

### *6.2.2 Heterosis Performance in Foxtail Millet*

Many studies have shown that hybrid vigor, or heterosis, of foxtail millet is evident in many characteristics including seedling viability, growth potential, biotic and abiotic stress resistance, and grain yield. However, the specific advantages differ widely among different hybrids.

Compared with their parent lines, hybrids showed faster seed germination and 1–2 days' earlier seedling emergence from soil (Cui et al. 1979), indicating hybrid growth vigor over conventional lines. The growth-related characterization of six foxtail millet hybrids, including '38A' × 'Hui 329,' showed that the dry weight of hybrid seedlings was 15 % greater than that of their parents. Also, the hybrid plants were taller, produced more secondary roots, and had a greater stem diameter, compared with their parents and the control. Li et al. (1963) reported that the heading date of hybrid foxtail millet is usually earlier than those of their parents. Cui et al. (1987) observed that the foxtail millet hybrids 'Suanxi 28' × 'Zhangnong 10' and 'Huangxi 4' × '1007' were more drought resistant than conventional varieties. A drought in 1984 resulted in reduced heading or no heading in conventional varieties, but did not affect heading in these three foxtail millet hybrids.

Many studies have indicated that heterosis is more evident for grain yield per plant than for other yield-related characters, with grain yield being 20–30 % higher in hybrids than in their parents (Li et al. 1963; Du 1984). The main reason for this increase is increased number of spikelets per panicle. However, not all hybrids show heterosis in grain yield. In one study, 57 % of 533 hybrids showed heterosis

over their parents, but only 4 % of those hybrids produced a higher grain yield than that of the local control cultivar (Cui et al. 1987). In addition to grain yield, some hybrids show heterosis in protein content, vitamin content, and other characteristics (Ji 1990).

The performance of foxtail millet heterosis is closely related to the geographical origin of their parents and/or genotype differences. Among 251 hybrids produced from the male sterile lines 'Suanxi 28' and 'Huangxi 4,' hybrid vigor was stronger in the offspring of distantly related parents than in the offspring of two local parents. Hybrids produced with cultivars as parents showed 71.4 % heterosis, while hybrids produced with landraces as parents showed only 30.0 % heterosis, indicating that the newly developed cultivars are advantageous in hybrid foxtail millet production (Cui et al. 1987).

### *6.2.3 Development of CMS Lines of Foxtail Millet*

In the early days of development of male sterile foxtail millet lines in the 1960s and 1970s, spontaneous male sterile mutations identified in farm fields were crossed with other genotypes to develop CMS lines, but no true CMS line was developed. Cui Wensheng from the Baxia Agricultural Institute and other breeders conducted many crosses with landraces from distant geographical origins, and although these crosses created several male sterile foxtail millet lines, no CMS line was developed (Cui and Gui 1984). As the wild ancestor of domesticated foxtail millet, green foxtail has been used as the maternal parent in crosses with domesticated foxtail millet to develop CMS lines. More than 30 crosses with green foxtail were conducted in 1987 by Zhu et al. and several hundred similar crosses were conducted the 1970s by different breeders in China. Although a few male sterile lines were created (Zhu et al. 1987; Zhi et al. 2007), no true CMS line was developed. In the late 1980s and early 1990s, the research group lead by Zhu Guanqin at the Shanxi Academy of Agricultural Sciences conducted many distant crosses between *S. faberii*, *S. verticillata,* and *S. yunnanensis* as the female parents and foxtail millet as the male parent. Hybrid plants were successfully obtained from *S. faberii* and *S. verticillata* as parents, but no $F_1$ individual was obtained from the cross using *S. yunnanensis* as the female parent, even when hybrid embryos were rescued by tissue culture (Zhou et al. 1988). Because of the tetraploid nature of *S. faberii*, *S. verticillata*, and *S. yunnanensis*, tetraploid foxtail millet lines created by artificial chromosome duplication were used for crosses, and then back-crossed several times with diploid foxtail millet to restore the diploid character. Among the two successful distant crosses with *S. faberii* and *S. verticillata* as the female parents, a single hybrid offspring of *S. verticillata* was developed into a CMS line. It was designated as the Ve CMS line of foxtail millet, because its cytoplasm comes from *S. verticillata* (Zhu et al. 1991). This is the only CMS line reported in foxtail millet breeding so far, and for unknown reasons this line has not been used to produce hybrid seeds.

### 6.2.4 *Identification of the Ch Dominant Male Sterile Line in Foxtail Millet*

Hu et al. (1986) reported the first genetic dominant male sterile line 'Ch78182' identified among the offspring of a cross between the Australian landrace Aodaliyagu, (meaning Australia) and the Chinese landrace Tulufan from Xinjiang. 'Ch78182' was designated as the Ch dominant male sterile line, because it was developed in Chifeng Agricultural Institute of Inner Mongolia. The individual male sterile plant was first identified in 1976, and an inheritance study confirmed its genetically dominant nature (Hu et al. 1993). Screening for restoration lines identified that '181–5' could fully restore the male fertility of 'Ch78182.' An inheritance study showed that lines '181–5' carry the epistasis gene *GG*, which suppresses the expression or function of the dominant male sterile gene *ChCh*. These two genes, *GG* and *ChCh*, were found to be linked in the foxtail millet genome. The male sterile individuals could be genotyped as ChChgg and Chchgg, and any genotype with --G- was male fertile (Hu et al. 1993). The main reason for the male sterility of 'Ch 78182' line was shown to be the elimination of anther dehiscence. That study showed that pollen development is nearly normal inside the anther, and fully developed mature pollen grains accumulate inside the anthers (Diao et al. 1991). Later, Hu et al. (1993) identified that 'Ch 78182' could self-pollinate to some degree, leading to 6–10 % seed setting in Hainan Province of Southern China, implying that environmental factors affect anther dehiscence. Plants from those self-pollinated seeds are thoroughly male sterile in northern China, where foxtail millet is cultivated for grain production. According to all those characters of the dominant male sterile line 'Ch 78182' and its putative restoration lines '181–5,' a hybrid seed production system was suggested, as shown in Fig. 6.1 (Hu et al. 1993). However, because of the complexity of this system, it has not been used for the commercial production of foxtail millet hybrid seeds.

### 6.2.5 *Development of Photo- or Thermo-Sensitive Genic Male Sterile Lines*

Photo- or thermo-sensitive genic male sterility (PTGMS) was successfully used in rice hybrid seed production in China from the 1990s, and this captured the interest of many foxtail millet breeders. The first foxtail millet PTGMS line '292' was developed in 1989 from pedigree selection of a cross between two landraces, Cai 5 and Ce35-1, at the Baxia Institute of Agricultural Sciences (Zhao et al. 1994). The '292' line was made male fertile to some degree by a short photoperiod treatment (13 h or shorter), and made male sterile by a longer photoperiod treatment (14 h or longer). This suggested that male sterile lines could be produced in short-day conditions (winter) in Hainan by self-pollination of '292,' and hybrid seeds could be produced in long-day conditions in summer in northern China. Based on line '292,' many crosses were made

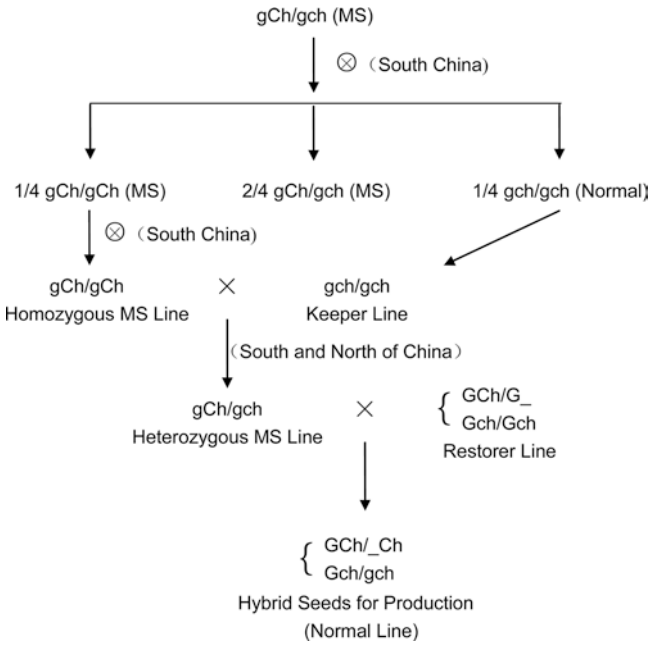


**Fig. 6.1** Chart of hybrid seed production using Ch dominant male line, modified from Hu et al. (1993). *Chch* represents dominant male sterile gene and *Gg* represents epistasis gene of *Ch*. Chchgg genotype can self-pollinate in southern China (Hainan or Zhanjiang) with a 6–10 % seed setting rate, thereby producing ChChgg and chchgg lines. ChChgg genotype can maintain itself in southern China by self-pollination, and is crossed with chchgg genotype to enlarge male sterile seed pool. Chchgg is crossed with GG restoration line to produce hybrid seeds for commercial foxtail millet production

with different landraces and cultivars, and different PTGMS lines were developed, including '821' which was much more stable than '292.' An inheritance study on '821' identified that the PTGMS male sterile gene was a single recessive one (Zhao et al. 1996). Wang et al. (2003) developed the PTGMS lines 'SMPA 1,' 'SMPA 2,' 'SMPA 3,' and 'SMPA 4,' all derivatives of '292.' The grain yields of the hybrids produced from those lines were more than 15 % higher than that of the local control cultivar. Another PTGMS foxtail millet line, 'Guang A1,' was developed by Cui et al. (1991). This line was derived from pedigree selection of the cross between the landraces Aodaliyagu and Zhongweizhuyeqing. 'Guang A4' was developed from the offspring of a cross between 'GuangA1' and the landrace Jiugenqi by Zhao and Cui (1994). Characterization of photosensitivity showed that a short photoperiod treatment (10–12 h) resulted in 45 % sterile pollen in 'Guang A4' with 33.2–38.2 % seed setting by self-pollination. A longer photoperiod treatment (15 h) resulted in 99 % sterile pollen with a 0.1 % seed setting rate (Zhao and Cui 1994). Although there have been several reports of PTGMS in foxtail millet, no hybrid cultivar developed using this technology has been released so far. This is probably because the stability of the male sterile character of PTGMS differs in unstable environmental conditions and from year to year. This should be studied in detail in the future.

### 6.2.6 *Successful Heterosis Utilization in Foxtail Millet with PAGMS Lines*

Cui Wenshen from the Baxia Institute of Agricultural Sciences identified an individual male sterile foxtail millet plant from the landrace Hongmiaosuanpibai in 1969. Consecutive selections developed this male sterile plant into a partial genic male sterile line (PAGMS) line, which was named 'Suanxi 28.' A characterization and inheritance study demonstrated that 97% of its spikelets were male sterile and 3% were male fertile. With its 3% male fertile spikelets, 'Suanxi 28' could achieve approximately 5% seed setting by self-pollination, and the progeny of those self-pollinated seeds retained the 97% male sterile character. An inheritance study of the male sterile character demonstrated that it is controlled by a single recessive gene, and all normal landraces or cultivars can restore male sterile to male fertile ones (Cui et al. 1979). This is a distinct type of male sterility, which was named Gaoduxiongxingbuyu in Chinese, meaning highly male sterile, but the term "partial genic male sterile line" is more appropriate in English (Diao 2014). Similar to 'Suanxi 28,' other PAGMS lines were also developed from different source of male sterile origin, such as 'Huangxi 1,' 'Huangxi 4,' and 'Huangmi 1.' The male sterile genes of 'Suanxi 28' and 'Huangxi 1' were transferred into different foxtail millet genotypes and several PAGMS lines were developed. Other PAGMS lines similar to 'Suanxi 28' were developed at different institutes in China in the late 1980s and 1990s, such as '1066 A,' '350 A,' 'Jinfen A,' and 'Gao 146A' (Wang et al. 1993; Li et al. 2011; Yang et al. 2007).

The advantage of this system is that it is easy to maintain male sterile seeds with one line and easy to screen restoration lines from common cultivars. The disadvantage of this system is that false hybrid seeds must be eliminated from the hybrid seed pool. In the late 1970s, false hybrid plants were removed by hand at the seedling stage in field trials, and were identified based on color, as false male sterile hybrid seedlings were yellow instead of green. The labor-intensive nature of this system restricted the use of these hybrid cultivars, even though several cultivars showed substantially increased grain yields, such as 'Suanxi 28' × 'Zhangnong 10' and 'Huangxi 4' × '1007' (Cui et al. 1979). In the first decade of this century, an herbicide resistance gene was transferred into foxtail millet from wild green foxtail, and herbicide-resistant cultivars were released (see Chap. 15). Hybrid seeds produced using herbicide-resistant foxtail millet cultivars as the restoration line retain the herbicide-resistant character, and so false male sterile plants can be easily killed by spraying with herbicide. This has created a new opportunity for the successful utilization of PAGMS foxtail millet lines (Wang et al. 1996).

The development of PAGMS lines and herbicide-resistant restoration lines has made it possible to utilize foxtail millet heterosis. Consequently, more than ten hybrid cultivars produced using this system have been released during the past 10 years in China. Most of these cultivars were developed at the Zhangjiakou Institute of Agricultural Sciences in Hebei Province. According to data from the National Test for Regional Adaptability (TRA) of foxtail millet cultivars at the National Extension Center for Agricultural Technology from 2000 to 2002, the grain yield of

hybrid foxtail millet 'Zhangzagu 1' was 5494.5 kg/ha, 19.78 % greater than that of the control variety 'Datong 14.' In the TRA trial in 2004, Zhangzagu 3 had a grain yield of 4455 kg/ha, 19.13 % greater than that of the control variety. In the TRA trial in Shanxi Province, the grain yield of the foxtail millet hybrid 'Changzagu 2' was 17 % greater than that of the control.

Today, foxtail millet is an important crop in arid and semi-arid regions, and it is potentially important for a much drier and warmer climate in the future. As a self-pollinated diploid species with a small genome, foxtail millet is fast becoming a new model plant for functional genomics and its wild progenitor, green foxtail, is becoming a model plant for research on $C_4$ photosynthesis and abiotic response. Progress in research on the germplasm, conventional breeding, functional genomics, and molecular breeding of foxtail millet will certainly improve research on heterosis in this crop.

Several breeding groups in China are focusing on the utilization of foxtail millet heterosis with PAGMS lines and have developed new super hybrid cultivars. However, few researchers are attempting to create new types of male sterile lines. Since 2007, three foxtail millet hybrid varieties have been released in Shanxi Province, and at least five new varieties are being evaluated in field trials in the National Test for Regional Adaptation of foxtail millet cultivars. Research efforts have improved foxtail millet heterosis utility, and analyses of population structure and heterotic groups by combining ability tests and molecular methods are underway. Therefore, we forecast a promising future for the utilization of foxtail millet heterosis.

## 6.3 Perspectives for Foxtail Millet Breeding

As a cereal with a long history of cultivation as a staple crop in arid and semiarid regions, foxtail millet is now recognized as a crop suitable for cultivation in the drier and warmer conditions predicted in the future. As such, it is receiving more attention as a sustainable agricultural crop (Li and Brutnell 2011; Lata et al. 2013; Diao et al. 2014). In the past 60 years in China, four great achievements have been made in foxtail millet breeding. First, a comprehensive germplasm collection of *Setaria* was established, comprising 27,760 accessions from all of the ecological regions of foxtail millet cultivation in China. These accessions are housed at the Chinese Gene Bank. Second, the grain yield and food quality of foxtail millet were greatly improved in the 1980s through the development of the super cultivars 'Yugu 1' and 'Zhaogu 1,' with most of the recently released cultivars being derivatives of these two genotypes. Third, the introduction of the gene conferring resistance to the herbicide sethoxydim from green foxtail into foxtail millet made its cultivation system more suitable for modern agriculture. Finally, the development of hybrid cultivars of foxtail millet through the PAGMS two-line system created a new method of heterosis utilization for this crop, distinct from the well-known CMS and PTGMS systems. All of those achievements have not only enriched the theory of plant breeding but have also established a firm foundation for the development of *Setaria* as a model species.

Compared with rice, wheat, maize, and other major cereals, conventional and hybrid cultivars of foxtail millet still have a low grain yield potential. Increasing the yield potential is one of the main goals of foxtail millet breeding. In northern China, foxtail millet is mainly consumed as a gruel for breakfast and dinner, mainly because the kernels of foxtail millet are more suitable for gruel than for preparation as a main-meal grain like rice. If the quality of foxtail millet kernels can be improved to a point where it can be prepared as a main food, then this will promote foxtail millet as a future main crop. Foxtail millet is a well-known drought-tolerant crop. This has been learned mainly from experience, and there is no detailed scientific data on its drought tolerance or water use efficiency. Consequently, it has been difficult to establish breeding programs to improve its drought tolerance and water use efficiency. Other characteristics that are important in breeding include resistance to diseases such as leaf rust and blast disease, and pests such as nematodes and other insects.

The recent advances in genome sequencing and functional genomics research on *Setaria* will lead to substantial changes in the breeding methods for foxtail millet in the near future. The elucidation of characters related to photosynthesis and grain yield and the molecular interaction networks between them will accelerate improvements in the yield potential of foxtail millet. It is possible that foxtail millet will become a major crop in dry land agriculture, with a grain yield potential similar to those of maize and sorghum. Using a combination of association and linkage mapping populations, the Foxtail Millet Research Team at the Chinese Academy of Agricultural Sciences (CAAS) is now elucidating coding genes and dominant alleles of foxtail millet related to taste and commercial and nutritional quality, with the aim to improve foxtail millet to the point where it can become a widely used crop. Breeders at the Institute of Millet Crops, Hebei Academy of Agricultural Sciences, and at the Institute of Crop Sciences, CAAS, are currently developing markers for resistance to leaf rust and blast disease. Several research groups in India, China, and America are currently deciphering genes, alleles, and molecular networks related to drought resistance of foxtail millet (Wang et al. 2016; Qie et al. 2014; Qi et al. 2013). The results of those studies, as well as recent progress in research on the drought resistance of other crop species, will be useful for the molecular breeding of highly drought-tolerant cultivars of foxtail millet and other crops in the near future.

## References

Ahanchede A, Poirier-Hamon S, Darmency H. Why no tetraploid cultivar of foxtail millet. Genet Resour Crop Evol. 2004;51(3):227–30.

Bennetzen JL, Schmutz J, Wang H, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30(6):555–61.

Chen X, Wu Q, Zhu J. Artificial creation of double diploid of *Setaria italica-Setaria faberlli* hybrid. J Shanxi Agric Sci. 1997;5:3–4.

Cheng R, Liu Z. Special hybridization technologies for foxtail millet. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 206–21.

Cheng R and Dong Z Breeding and production of foxtail millet in China. In Cereals in China, 2010, Limagrain and CIMMYT.
Cui W, Gui D. Study on foxtail millet male sterile materials from Zhangjiakou and their utility. Acta Agric Bor Sin. 1984;9:4–9.
Cui W, Ma H, Zhang D. The selection and utilization of "Sun Hsi 28"—a male-sterility strain of millet. Sci Agric Sin. 1979;12(1):43–6.
Cui WS, Du G, Kong YZ. The selection and utilization of two-line hybrid of millet. In: Millet Crop Section of Chinese Association of Crop Sciences, editor. Synopses of millets in China. Shijiazhuang: Millet Crop Section of Chinese Association of Crop Sciences; 1987. p. 80–90.
Cui W, Kong Y, Zhao Z, et al. A preliminary observation on breeding of photoperiod-sensitive and recessive GMS material of foxtail millet "292". Acta Agric Bor Sin. 1991;S1:47–52.
Diao X. Foxtail millet production and research system in China. Beijing: China Agricultural Science and Technological Press; 2011.
Diao X. Problems and development perspectives on heterosis utility in foxtail millet. In: Cai J, editor. Utilization of heterosis in crops. Beijing: Higher Education Press; 2014. p. 237–43.
Diao X, Du R, Wang T, et al. Cytomorphological study on anther development of Ch dominant GMS plants in foxtail millet. Acta Agric Bor Sin. 1991;1:13–7.
Diao XM, Wang P, Zhi H. Somaclonal variation of foxtail millet and its application in breeding. Acta Agron Sin. 2002;28(4):480–5.
Diao XM, Li H, Zhi H. Genetic improvement of foxtail millet. In: Lu Q, Zhao T, editors. Genetic improvement of crops. Beijing: The Chinese Agricultural Scientific and Technology Press; 2011. p. 546–90.
Diao X, Schnable J, Bennetzen JL, Li J. Initiation of vSetaria as a model plant. Front Agr Sci Eng. 2014;1(1):16–20.
Du S. Heterosis in $F_1$ hybrids and their relationship with parents. In: Heilongjiang Academy of Agricultural Sciences, editor. A collection of foxtail millet scientific research in Northeastern China. Harbin: Heilongjiang Academy of Agricultural Sciences; 1984. p. 108–11.
Gao J, Wang R. Foxtail millet grain quality improvement. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 206–21.
Hu H, Ma S, Shi Y. The discovery of a dominant male-sterile gene in millet (*Setaria italiica*). Acta Agron Sin. 1986;12(2):73–8.
Hu H, Shi Y, Wang C, et al. Studies on the inheritance of "Ch-" dominant nucleus sterility of millet (*Setaria italica*) and its applications to commercial production. Acta Agron Sin. 1993;19(3): 208–17.
Ji G. Gene and environmental effects on protein contain of foxtail millet. Millet Crops. 1990;1:13–5.
Jia G, Huang X, Zhi H, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013;45(8):957–61.
Lata C, Gupta S, Prasad M. Foxtail millet: a model crop for genetic and genomic studies in bioenergy grasses. Crit Rev Biotechnol. 2013;33(3):328–43.
Li Y. Foxtail millet breeding. Beijing: China Agricultural Press; 1997.
Li P, Brutnell TP. *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. J Exp Bot. 2011;62(9):3031–7.
Li CH, Pao WK, Li HW. Interspecific crosses in Setaria. II. Cytological studies of interspecific hybrids involving 1, *S. faberii* and *S. italica*, and 2, a three way cross, F2 of *S. italica* × *S.viridis* and *S. faberii*. J Hered. 1942;33:351–5.
Li HK, Jia KF, Li H. A study of foxtail millet variation and heterosis of $F_1$ hybrid. In: Heilongjiang Academy of Agricultural Sciences, editor. A collection of foxtail millet scientific research in Northeastern China. Harbin: Heilongjiang Academy of Agricultural Sciences; 1963. p. 40–50.
Li H, Wang Y, Tian G. Practice and thinking on millet heterosis utilization in Shanxi. J Shanxi Agric Sci. 2011;39(10):1035–9.
Li W, Zhi H, Zhang S, et al. Morphological effect and genomic mapping of Si-SP1 (small panicle) in foxtail millet. J Plant Genet Resour. 2015;16(3):581–7.
Li W, Tang S, Zhang S, et al. Gene mapping and functional analysis of the novel leaf color gene *SiYGL1* in foxtail millet [*Setaria italica* (L.) P. Beauv]. Physiol Plant. 2016;157(1):24–37.

Luo X, Guo F, Zhou J, et al. Immature embryo culture of a cross between *Setaria italica* (Ch4n) and *S. faberii* and studies on the morphological and cytological characteristics of the F1 plant. Acta Agron Sin. 1993;19(4):352–8.

Qi X, Xie S, Liu Y, Yi F, Yu J. Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. Plant Mol Biol. 2013;83:459–73.

Qie L, Jia G, Zhang W, Schnable J, Shang Z, Li W, Liu B, Li M, Chai Y, Zhi H, Diao X. Mapping of quantitative trait locus (QTLs) that contribute to germination and early seedling drought tolerance in the interspecific cross *Setaria italica* × *Setaria viridis*. PLoS One. 2014;9(7):e101868.

Takahashi N. Studies on the flower of Italian millet and a method of its artificial hybridization (in Japanese). Proc Crop Sci Soc Jpn. 1942;13:337–40.

Wang DY, Guo GL. Ecological regions and ecotypes of Chinese foxtail millet. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 44–77.

Wang T, Du R, Hao F. Studies and utilization of highly male sterility of summer millet (*Setaria italica* B.). Sci Agric Sin. 1993;26(6):88.

Wang RQ, Gao JH, Wang ZX, Wang ZM. Establishment of trisomic series of millet (*Setaria italica* [L.]Beauv.). Acta Bot Sin. 1994;36(9):690–5.

Wang T, Du R, Chen H, et al. A new way of using herbicide resistant gene on hybrid utilization in foxtail millet. Sci Agric Sin. 1996;29(4):96.

Wang T, Zhao Z, Yan H, et al. Gene flow from cultivated herbicide-resistant foxtail millet to its wild relatives: a basis for risk assessment of the release of transgenic millet. Acta Agron Sin. 2001;27(6):681–7.

Wang Y, Wang S, Li H, et al. Studies on breeding of photoperiod sensitive male-sterile line in millet and its application. Sci Agric Sin. 2003;36(6):714–7.

Wang C, Jia G, Zhi H, et al. Genetic diversity and population structure of Chinese foxtail millet [*Setaria italica* (L.) Beauv.] landraces. G3 Genes Genom Genet. 2012;2(7):769–77.

Wang Y, Li L, Tang S, Liu J, Zhang H, Zhi H, Jia G, Diao X. Combined small RNA and degradome sequencing to identify miRNAs and their targets in response to drought in foxtail millet. BMC Genet. 2016;17:57.

Wu Q, Bai J. Cytogenetic and isoenzymic studies on Setaria millet and *S. verticillata* (2X) and *S. verticiformis* (4X). Acta Botan Boreali Occiden Sin. 2000;20(6):954–9.

Yang C, Chen Y, Shi G, Hou G, Shi G, Shi H. Development of foxtail millet male sterile line Jinfen 1A. J Shanxi Agric Sci. 2007;6:1–3.

Yi HY. Radiation and chemical induced breeding in foxtail millet. In: Li Y, editor. Foxtail millet breeding. Beijing: China Agricultural Press; 1997. p. 260–88.

Zhang W, Zhi H, Liu B, Wang X, Pang Z, Li J, Wang G, Li M, et al. Phenotype variation and vertical distribution of foxtail millet root system in RIL from a cross of Yugu 1 and wild green foxtail W53. Acta Agron Sin. 2014;40(10):1717–24.

Zhao Z, Cui W. PTGMS development from crosses between different geographic origin of foxtail millet landraces. Agric Sci Bull. 1994;5:24–5.

Zhao Z, Cui W, Du G. Primary study on response of photoperiod sensitive nuclear genetic male sterile foxtail millet to sunshine. Acta Agric Bor Sin. 1994;S1:29–32.

Zhao Z, Cui W, Du G, et al. The selection of millet photo (thermo) sensitive sterile line 821 and a study on the relation of sterility to illumination and temperature. Sci Agric Sin. 1996; 29(5):23–31.

Zhi H, Wang Y, Li W, et al. Development of CMS material from intra-subspecies hybridization between green foxtail and foxtail millet. J Plant Genet Resour. 2007;8(3):261–4.

Zhou J, Luo X, Guo F, et al. Plant regeneration in tissue culture of Setaria yunnanensis × *S. italica* (4n) F1 plants. Acta Agric Sin. 1988;14:227–31.

Zhu G, Li X, Shi G, Li Y. Autotetraploid of foxtail millet sterile lines introduced by Colchicine. J Shanxi Agric Sci. 1987;6:33.

Zhu G, Wu Q, Ma Y. The development of Ve type male sterile line of foxtail millet. J Shanxi Agric Sci. 1991;1:7.

# Chapter 7
# Genetic Differentiation and Crop Evolution of Foxtail Millet

**Kenji Fukunaga**

**Abstract** Several studies on genetic differentiation and crop evolution based on intraspecific hybrid pollen semi-sterility, isozymes, ribosomal DNA (rDNA) RFLP, nuclear RFLP, mitochondrial DNA (mtDNA) RFLP, RAPD, AFLP, transposon display (TD) markers, and single nucleotide polymorphisms (SNPs) have been carried out to elucidate genetic relationships of foxtail millet accessions, mainly from Eurasia. Most of the studies suggest that China is the center of diversity of foxtail millet and that landraces are grouped in geographical groups. Evolution of two genes, *waxy* gene controlling amylose content in endosperm and *polyphenol oxidase* (*PPO*) gene for phenol color reaction (Phr) in grains, was also reviewed. These analyses showed that multiple independent loss-of-function mutations occurred in each of the two genes under human/natural selection.

**Keywords** Center of diversity • Crop evolution • Domestication • DNA markers • Geographical variation • Genetic differentiation • Phylogeny • *Polyphenol oxidase* (*PPO*) gene • *Waxy* gene • Foxtail millet • Green foxtail

## 7.1 Hypotheses on Geographical Origin of Foxtail Millet and Crop Evolution of Foxtail Millet

Foxtail millet, *Setaria italica* (L.) P. Beauv., is one the oldest domesticated cereals in the Old World. Archaeological remains of foxtail millet were found in sites of Peiligang and Cishan near the Yellow River, dating back to ca. 5000–6000 BC (Li and Wu 1996), in prehistoric sites in Europe (Kűster 1984) and in Transcaucasus (Lisitsina 1976). Foxtail millet has been utilized in various ways peculiar to each area of Eurasia (Sakamoto 1987), and it is thought to have played an important role in early agriculture in the Old World.

The geographical origin of foxtail millet is still a controversial issue. Cytological studies indicate that the wild ancestor of foxtail millet is green foxtail (*S. italica* ssp.

K. Fukunaga (✉)
Prefectural University of Hiroshima, Shobara 727-0023, Japan
e-mail: fukunaga@pu-hiroshima.ac.jp



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_7

*viridis* = *S. viridis*) (Kihara and Kishimoto 1942; Li et al. 1945). However, the geographical origin of domesticated foxtail millet cannot be determined from the distribution of ssp. *viridis*, since it is found commonly in various areas in Europe and Asia (and also currently in the New World). Vavilov (1926) stated that the principal center of diversity for foxtail millet is East Asia, including China and Japan. Harlan (1975) suggested independent domestication in China and Europe based on archaeological evidence. The archaeological, isozyme, and morphological evidences (de Wet et al. 1979; Jusuf and Pernes 1985; Li et al. 1995a, b) suggest that China is the center of diversity and origin of foxtail millet but independent origins in other regions of this millet cannot be excluded. Further, Li et al. (1995b) stated that landraces in Afghanistan and Lebanon had been domesticated independently in relatively recent times because they had primitive morphological characters such as several tillers with small panicles and look like ssp.*viridis* but have non-shattering large grains. Molecular analyses support the view that foxtail millet landraces have differentiated into local landrace groups and that China is the center of diversity (isozymes: Jusuf and Pernes (1985); prolamin: Nakayama et al. (1999); rDNA: Fukunaga et al. (1997a, 2006), Schontz and Rether (1998), Eda et al. (2013); RAPD: Li et al. (1998), Schontz and Rether (1999); AFLP: Le Thierry d'Ennequin et al. (2000); genomic RFLP: Fukunaga et al. 2002b, mitochondrial DNA RFLP: Fukunaga and Kato 2003, transposon display; Hirano et al. 2011). Recent archaeological evidence also supports the domestication of foxtail millet in China (Nasu et al. 2007; Hunt et al. 2008). In contrast to the hypothesis of Chinese origin and multiple origin, Sakamoto (1987) suggested that foxtail millet originated somewhere in Central Asia–Afghanistan–Pakistan–India because accessions with less compatibility (Kawase and Sakamoto 1987) and with primitive morphological traits are found there. This hypothesis, which excludes China as a center of origin of foxtail millet, is very different from the others.

In this chapter, we review studies on genetic differentiation of foxtail millet landraces from various parts of Europe and Asia in morphological/agronomic characters, biochemical markers, intraspecific pollen sterility, and DNA markers. We also review evolution of two genes, *Waxy* and *Polyphenol oxidase* (*PPO*), which were selected during domestication and dispersal of foxtail millet.

## 7.2 Variation in Morphological Characters

Variation in morphological and agronomic characters involved in domestication and diversification of foxtail millet has been described and analyzed: (Dekaprelevich and Kasparian 1928; de Wet et al. 1979; Kawase 1986; Takei and Sakamoto 1987, 1989; Prasada Rao et al. 1987; Ochiai et al. 1994; Li et al. 1995a, b; Ochiai 1996; Fukunaga et al. 1997b; Hammer and Khoshbakht 2007). Some researchers classify foxtail millet landraces into two to four subspecies/races such as moharia, maxima, indica, and nana (de Wet et al. 1979; Prasada Rao et al. 1987; Li et al. 1995b), but criteria for the classification are ambiguous. Kawase (1986) and Ochiai et al. (1994) investigated

variation in morphological/agronomic characters such as plant height, number of tillers, panicle length, and number of days to heading of foxtail millet landraces from Europe and Asia. Nguyen and Pernes (1985) and Li et al. (1995b) also investigated variation of morphological/agronomic characters of foxtail millet landraces by multivariate analyses. These works indicate that morphologically primitive landraces characterized by several tillers with small panicles, which look like ssp.*viridis* but have non-shattering large grains, are distributed in Afghanistan, northwestern Pakistan, Central Asia, and Lebanon whereas most of accessions from other regions such as East Asia have no or a few tillers with large panicle(s). Hammer and Khoshbakht (2007) also reported cultivation of morphologically primitive landraces in Northern Iran. A few researchers insist that foxtail millet was domesticated in Central Asia–Pakistan–Afghanistan–northwest India because the morphologically primitive type was cultivated there (Sakamoto 1987; Ochiai 1996), whereas Li et al. (1995b) insist that the morphologically primitive type with several tillers and small panicles, which look like ssp. *viridis,* was domesticated independently. Description and analyses of morphological variation are important but not sufficient to address questions on geographical origin(s) and phylogeny of foxtail millet landraces.

## 7.3 Genetic Differentiation of Foxtail Millet Landraces, Revealed by Biochemical and Genetic Markers and Intraspecific Hybrid Pollen Sterility

Several studies have been carried out to clarify the genetic relationships of foxtail millet from Europe and Asia (and partly from Africa) such as (1) biochemical markers (isozymes and prolamin), (2) intraspecific hybrid pollen sterility, and (3) DNA markers (nuclear RFLP, mitochondrial RFLP, RAPD, AFLP, transposon display (TD) markers, and single nucleotide polymorphisms (SNPs)). As summarized in Table 7.1, these studies revealed that foxtail millet landraces differentiated into local geographical groups and that East Asian landraces (in particular Chinese landraces) are the most diverse.

These genetic studies are described as follows:

1. Variation in biochemical markers (isozymes and prolamin)

   Kawase and Sakamoto (1984) investigated polymorphism in two loci, *Est-1* and *Est-2* of the esterase isozymes of 432 accessions of foxtail millet collected from different areas throughout Eurasia by gel isoelectric focusing. On locus *Est-1*, most of accessions had *Est-1 a*, which was widely distributed throughout Eurasia, 9% of accessions had *Est-1 b* which was distributed in China and Korea. On locus *Est-2*, most of accessions had *Est-2 a*, but nine had *Est-2 b*, which is found in all of the accessions from the western part of Europe and in one of the Indian accessions. Six had the *Est-2 c* allele, which was found in Japan and China. They concluded that the distribution of *Est-2 a* and *-2 b* might indicate some degree of phylogenetic differentiation between the Asian and the European accessions and that Chinese accessions showed polymorphisms in both loci.

**Table 7.1** Genetic studies on genetic differentiation of foxtail millet landraces and geographical groups and centers of diversity

| Genetic markers/ intraspecific hybridity/ pollen sterility | Geographical groups | Center of diversity | References |
|---|---|---|---|
| Esterase isozymes | East Asia vs. Europe | East Asia | Kawase and Sakamoto (1984) |
| Ten isozymes | China–Korea–Japan, Okinawa (Nansei Islands of Japan)–Taiwan, India–Kenya, Europe | | Jusuf and Pernes (1985) |
| Prolamine | Europe, Tropical Groups | China | Nakayama et al. (1999) |
| Hybrid sterility | China–Korea–Japan, Okinawa (Nansei Islands of Japan)–Taiwan, Lan-Hsű-Batan Islands India–Afghanistan, Europe | | Kawase and Sakamoto (1987), Kawase et al. (1997) and Kawase and Fukunaga (1999) |
| rDNA | China–Korea–Japan, Okinawa (Nansei Islands of Japan)–Taiwan–the Philippines, India, Afghanistan–Northern Pakistan | China | Fukunaga et al. (1997a, 2006, 2011) and Eda et al.(2013) |
| Nuclear RFLP | East Asia, Nansei Islands–Taiwan–the Philippines, India, Afghanistan–Central Asia–Europe | China | Fukunaga et al. (2002b) |
| mtDNA | Not clear | China | Fukunaga and Kato (2003) |
| RAPD | Central Europe and two Asiatic groups (north and south) | | Schontz and Rether (1999) |
| AFLP | Not clear | China | Le Thierry d'Ennequin et al. (2000) |
| TD | East Asia, Nansei Islands–Taiwan–the Philippines, India, Central Asia, Europe | China | Hirano et al. (2011) |
| SNPs | North China, Central–South China, South Asia, Central Asia, Europe, etc. | China | Jia et al. (2013) |

Jusuf and Pernes (1985) investigated the genetic diversity of a world collection of foxtail millet accessions and some samples of wild populations (ssp. *viridis*) by means of electrophoresis on five enzymes (ten loci) *Est*, *Acph*, *Got*, *Mdh*, and *Pgd*. They found some genetic groups of foxtail millet in China–Korea–Japan, Okinawa (Nansei Islands of Japan)–Taiwan, India–Kenya, and Europe. They also investigated wild populations collected in France and China and concluded that it is possible that there were independent domestications in both Europe and China because foxtail millet and *S. viridis* accessions were

more closely related in isozyme alleles in each of Europe and China. Nakayama et al. (1999) investigated allelic variation at the two prolamin loci (*Pro1* and *Pro2*) and their geographical distribution in 560 local cultivars of foxtail millet collected mainly from Eurasia and studied using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Two alleles (*Pro1a* and *Pro1null*) at the *Pro1* locus and six alleles (*Pro2a*, *Pro2b*, *Pro2c*, *Pro2d*, *Pro2e*, and *Pro2f*) at the *Pro2* locus were detected among the cultivars examined. No apparent trend in *Pro1* was observed in geographical distribution. In contrast, two common alleles at the *Pro2* locus, *Pro2b* and *Pro2f*, had clear differential geographical distribution. The *Pro2b* allele was most frequent in Europe and decreased in frequency eastwards. The *Pro2f* allele occurred frequently in subtropical and tropical regions including the Nansei islands of Japan, the Philippines, Nepal, India, Pakistan, and Africa. All eight alleles at the *Pro1* and *Pro2* loci occurred in China, suggesting China is a center of diversity. They also found a "tropical group" characterized by the *Pro2f* allele and other genes.

2. Classification by mean of intraspecific hybrid pollen sterility

   Intraspecific hybrid pollen sterility can be a genetic indicator of differences. In rice, classification based on hybrid sterility was carried out (Kato et al. 1928), and Asian rice varieties were classified into two main groups, japonica, and indica types. Kawase and Sakamoto (1987) crossed 83 accessions of *Setaria italica* collected from various areas throughout Eurasia with three tester strains from Japan (tester A), Lan Hsü Island of Taiwan (B) and Belgium (C). The accessions could be clearly classified into six types, designated as types A, B, C, AC, BC, and X. They regarded pollen fertility of more than 75 % as normal. The accessions of types A, B, and C were those that produced F1 hybrids having normal pollen fertility when crossed with testers A, B, and C, respectively. When both F1 hybrids from the crosses with two testers, A and C, or B and C, showed normal pollen fertility, the accession was classified as type AC or BC. The accessions whose F1 hybrids always showed pollen fertility of less than 75 % in all three cross combinations were designated as type X. Kawase et al. (1997) further investigated collections from northern Pakistan. Kawase and Fukunaga (1999) also used a landrace from Lan Hsü Island of Taiwan, which was classified into type X in Kawase and Sakamoto (1987), as tester D, and crossed it with landraces and reclassified type X in Kawase and Sakamoto (1987). Geographical distribution of these different types is shown in Fig. 7.1a. Most type A accessions were distributed in East Asia including Japan, Korea, and China. Type B accessions were narrowly found in Taiwan and in the southwestern part of the Nansei Islands of Japan. Most European accessions were found to be type C. Type D is distributed in Lan Hsü Island of Taiwan and Batan Islands of the Philippines. Accessions of types AC and BC were distributed in Afghanistan and India, respectively. Kawase and Sakamoto (1987) and Sakamoto (1987) concluded that Types AC and BC were less specialized genetically than A, B, or C and that the geographical distribution of these landrace groups suggests that *S. italica* was first domesticated in an area ranging from Afghanistan to India, and then dispersed both eastward and westward from there.

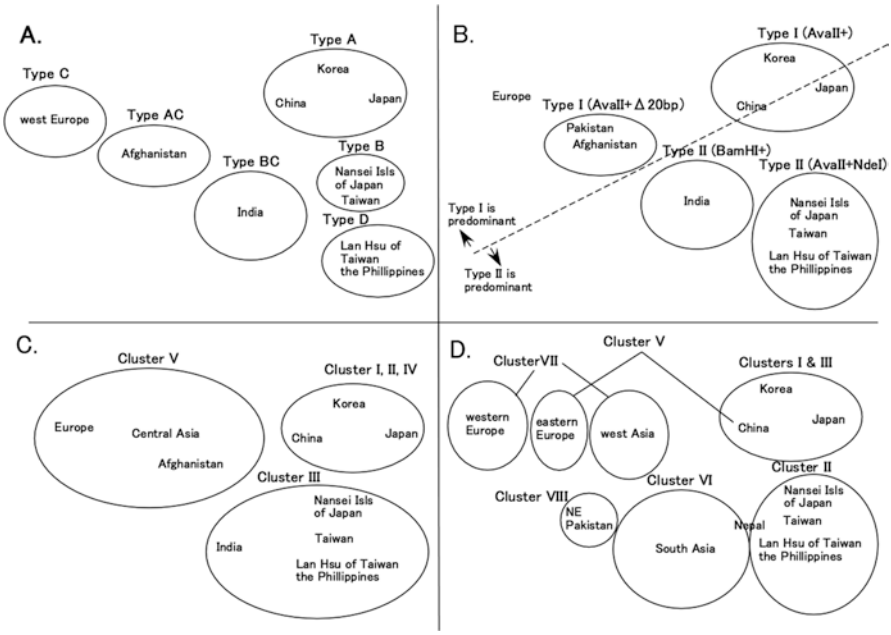


**Fig. 7.1** Schematic drawings of geographical distribution of different phylogenetic groups of foxtail millet based on four different analyses such as intraspecific hybrid pollen semi-sterility, rDNA PCR-RFLP, nuclear RFLP markers, and transposon diplay (TD markers). (**a**) Geographical distribution of intraspecific hybrid pollen semi-sterility types (Kawase and Sakamoto 1987; Kawase and Fukunaga 1999). (**b**) Geographical distribution of specific rDNA PCR-RFLP types (Eda et al. 2013). (**c**) Geographical distribution of clusters of a dendrogram showing genetic similarity between 62 accessions of foxtail millet based on 16 nuclear RFLP markers (Fukunaga et al. 2002b). (**d**) Geographical distribution of clusters of a dendrogram based on transposon display (TD) markers (Hirano et al. 2011)

3. DNA markers

(a) Ribosomal DNA (rDNA)

Fukunaga et al. (1997a) investigated restriction fragment length polymorphism (RFLP) and the structure of ribosomal RNA genes (rDNA) in 117 landraces of foxtail millet. Five RFLP phenotypes were found when the genomic DNA was digested with *Bam*HI; these were named types I–V. Of these types I, II, and III were the most frequent. Type I was mainly distributed in the temperate zone, type II in the Taiwan–Philippines Islands, and type III in South Asia. Restriction mapping of the cloned rDNA and comparison with RFLP phenotypes showed that the different types originated from a length polymorphism within the intergenic spacer (IGS) and *Bam*HI site changes within the IGS. Schontz and Rether (1998) also investigated rDNA in a world collection (43 accessions) for variation in repeat unit length and restriction enzyme site variability. They detected two lengths of repeat units of about 7.9 or 7.6 kb; the central European accessions and

most western European accessions have only the 7.6 kb repeat unit and most Asiatic lines have a 7.9 kb repeat unit, while lines originating from the north or the south of Asia showed different numbers of *Bam* HI fragments. These types correspond to types I–III in Fukunaga et al. (1997a). They concluded that the fact that difference between the Asiatic and European pool is not continuous (7.9 or 7.6 kb) excludes the hypothesis of domestication being based on the spread of an initial population over Eurasia. Fukunaga et al. (1997a) suggested that foxtail millet landraces differentiated into mainly two geographical groups, 7.6 kb repeat unit (=type I) from the temperate region and 7.9 kb repeat unit (=types II and III) from the subtropical–tropical region, whereas Schontz and Rether (1998) insisted that foxtail millet landraces differentiated into 7.9 kb repeat unit from Asia and 7.6 kb repeat unit from Europe. The difference in conclusions between these two studies are likely due to difference in number of Asian accessions used. Fukunaga et al. (2005) also determined the sequence of ribosomal DNA (rDNA) intergenic spacer (IGS) of foxtail millet isolated in the previous study and identified subrepeats in the polymorphic region. Fukunaga et al. (2006) sequenced ribosomal DNA intergenic spacer subrepeats and their flanking regions of foxtail millet landraces from various regions in Europe and Asia, as well as its wild ancestor green foxtail, to elucidate phylogenetic differentiation within each of types I–III found in the previous work and to elucidate relationships among these three types. Type I was classified into seven subtypes designated as Ia–Ig based on subrepeat sequences; C repeats downstream of those subrepeats were also polymorphic. Of these, subtypes Ia–Id and Ig were found in foxtail millet landraces. Subtypes Ia and Ib were distributed broadly throughout Asia and Europe. Subtype Ic was distributed in China, Korea, and Japan. Subtype Id has a 20-bp deletion in subrepeat 3 and has a unique C repeat sequence. This subtype was found in a morphologically primitive landrace group from Afghanistan and northwestern Pakistan and differed greatly from other type I subtypes, implying that these landraces were domesticated independently. Subtypes Ig was found in a landrace from Pakistan and Ia and Ie–Ig were in six wild ancestor accessions. Type II was also highly polymorphic and four subtypes were found and designated as subtypes IIa–IId, but sequence analyses indicated type III as monomorphic. This work indicates that type III should be classified as a subtype of type II (subtype IIe). Sequence polymorphism of subrepeats of types I–III indicated that subrepeats of subtype IIa are greatly divergent from others. Relationships among types I–III were much more complicated than anticipated based on previous RFLP work.

Ribosomal DNA spacer length polymorphisms were also studied in foxtail millet landraces from Pakistan and Afghanistan and in its wild ancestor (*S. viridis*) from Pakistan by PCR-based methods (Fukunaga et al. 2011). Sequence polymorphisms were investigated for accessions selected based on the observed length polymorphism. The PCR-based length and sequence polymorphisms of rDNA IGS clearly demonstrated genetic differentiation

between cultivated and wild forms in the region. Genetic differentiation was observed between different areas to some extent in the cultivated form and between different regions in the wild form from northern Pakistan. Recently, the rDNA PCR–RFLP of foxtail millet germplasm (480 accessions) collected throughout Eurasia and from a part of Africa was investigated with five restriction enzymes (Eda et al. 2013). Foxtail millet germplasm was classified by length of the rDNA IGS and RFLP, and clear geographical differentiation was observed between East Asia, the Nansei Islands of Japan–Taiwan–the Philippines area, South Asia, and Afghanistan–Pakistan (Fig. 7.1b). Evidence of migration of foxtail millet landraces between the areas was also found. Diversity indices (D) for each region were calculated, and it was concluded that the center of diversity of this millet is East Asia, including China, Korea, and Japan.

(b) Random Amplified Polymorphic DNA (RAPD) markers

Li et al. (1998) investigated random amplified polymorphic DNA (RAPD) in foxtail millet and wild relatives and confirmed that foxtail millet had been domesticated from *S. viridis*. Schontz and Rether (1999) investigated RAPDs in 37 accessions of cultivated *Setaria italica*, representative of Eurasian accessions. By using four 10-mer primers, they obtained 25 polymorphic bands and identified 33 different genotypes. A factorial analysis of correspondence was performed on the presence–absence data and three genetic groups were identified. These genetic groups were closely related to the geographical origin of the different accessions: one central European and two Asiatic groups (the first Asiatic accessions originating in latitudes below 35°N and the second comprising the Asiatic accessions originating in latitudes above 35° N).

(c) Nuclear genomic RFLP

Fukunaga et al. (2002b) investigated 16 RFLP loci in 62 landraces to study genetic differentiation in foxtail millet. Among 52 bands, 47 were polymorphic among foxtail millet landraces. A dendrogram based on RFLPs was divided into five major clusters (cluster I–V: Fig. 7.1c). Clusters I and II contained mainly accessions from East Asia. Cluster III consisted of accessions from subtropical and tropical regions in Asia, such as Nansei Islands of Japan, Taiwan, the Philippines, and India, and cluster IV consisted of some accessions from East Asia, an accession from Nepal and an accession from Myanmar. Cluster V contained accessions from central and western regions of Eurasia such as Afghanistan, Central Asia, and Europe. Chinese landraces were classified into four clusters. These results indicate that foxtail millet landraces have differentiated genetically between different regions and that Chinese landraces are highly variable.

(d) Mitochondrial DNA (mtDNA)

Mitochondrial DNA (mtDNA) was characterized by RFLPs in 94 accessions (Fukunaga and Kato 2003). Three RFLP patterns were observed by using rice *atp6* as a probe and were designated as types I–III. Difference between types I and II seem to be due to recombination between two *atp6* genes. In East and Southeast Asia and Afghanistan, both types I and II were found, while type I was predominant in India, Central Asia and Europe. In

China, type III was also found. Chinese accessions showed higher gene diversity than those from other regions. This result supported the previous studies on isozymes and nuclear RFLPs.

(e) AFLP

Le Thierry d'Ennequin et al. (2000) investigated AFLP markers to assess genetic diversity and patterns of geographic variation among 39 accessions of foxtail millet and 22 accessions of its wild progenitor. A high level of polymorphism was observed. Dendrograms based on Nei and Li distances from a neighbor joining procedure were constructed using 160 polymorphic bands. In contrast to other molecular marker studies, no specific geographic structure could be extracted from the data. The high level of diversity among Chinese accessions was consistent with the hypothesis of a center of domestication in China.

(f) Transposon Display (TD) markers

Hirano et al. (2011) investigated genetic structure by transposon display (TD) using 425 accessions of foxtail millet and 12 of the wild ancestor green foxtail. They used three recently active transposons (*TSI-1*, *TSI-7*, and *TSI-10*) as genome-wide markers and succeeded in demonstrating geographical structure for foxtail millet. A neighbor-joining dendrogram based on TD grouped the foxtail millet accessions into eight major clusters, each of which consisted of accessions collected from adjacent geographical areas (Fig. 7.1d). Eleven out of 12 green foxtail accessions were grouped separately from the clusters of foxtail millet. These results indicated strong regional differentiations and a long history of cultivation in each region. They also suggest a monophyletic origin of foxtail millet domestication.

(g) Single nucleotide polymorphisms (SNPs)

Recently, a large-scale analysis of whole genome single nucleotide polymorphism (SNP) in 916 accessions (mainly from China but also including accessions from other regions such as Japan and Korea, Southeast Asia, South Asia, Central Asia, Europe, Africa, and the USA) was carried out by Jia et al. (2013). They found that the 916 varieties can be clearly classified into two divergent groups, spring-sown form (type 1) with 292 varieties and summer sown form (type 2) with 624 varieties, and there was a clear geographical distribution of these two groups in the Chinese accessions—the majority of type 1 accessions were from northern China and high altitude areas of northwest China, whereas most of the type 2 accessions originated from central and southern China, which have warmer climates (see Chap. 2). As for accessions from other countries, they found that varieties from the same geographical regions tended to belong to the same clades in phylogenetic trees. They concluded that foxtail millet may have a single origin of domestication but that a deep investigation of the wild ancestor is needed. Their results that foxtail millet accessions can be divided into northern and southern groups may correspond with distribution of rDNA types I and II and results of other genetic studies although most of the materials that Jia et al. (2013) used were from China. Further analysis using more accessions from other countries will be also required.

## 7.4 Evolution of Two Genes (*Waxy* and *PPO*) Under Human/Natural Selection

In cereals, several genes involved in domestication and diversification have been studied in rice (e.g., Konishi et al. 2006; Ishii et al. 2013) and maize (e.g., Doebley et al. 1997; Wang et al. 2005) and six row and naked grains in barley (Komatsuda et al. 2007; Taketa et al. 2008). In foxtail millet, the *waxy* gene controlling amylose content in endosperm and the *polyphenol oxidase* (*PPO*) gene for phenol color reaction (Phr) in grains have been investigated, as it is known that these genes have evolved under human/natural selection in other cereals.

### *7.4.1 Evolution of Waxy*

Endosperm starch of cereals consists of amylose and amylopectin. Wild type (non-waxy) of endosperm starch consists of ca. 20 % or more of amylose and ca. 80 % of amylopectin, whereas the waxy type consists of ca. 100 % amylopectin and lacks amylose. The non-waxy type (*Wx*) is genetically dominant to the waxy type (*wx*). Endosperm starch with the recessive genotype, waxy type, has a stickier texture than the normally dominant non-waxy type. Both of these endosperm types are found among landraces of sorghum, rice, foxtail millet, maize, common millet, barley, and Job's tears (Sakamoto 1996). The waxy types of these cereals are found in East and Southeast Asia, but are rare in India and further westward. A core area where people show a strong ethnobotanical preference for waxy cereals, which extends from Southern China through Northern Thailand to Assam, has been identified (Sakamoto 1996; Yoshida 2002). In adjacent areas like Taiwan, Japan, and Korea, waxy cereals are grown mainly on upland soils, and are used in traditional rituals, or eaten only on special occasions. This trait is apparently associated with ethnological preferences in the areas (e.g., Fogg 1983; Takei 1994).

Waxy endosperm arises through the disrupted expression or loss of function of the *waxy* (*GBSS1*) gene that encodes granule-bound starch synthase I (GBSS I) (Sano 1984). Waxy-type cereals are characterized by little or no starch amylose, which constitutes about 20 % or more of the total starch in the non-waxy endosperm. This character has often been neglected in other regions, although waxy maize, which was first reported (Collins 1909) in Chinese landraces, is now globally used for the production of waxy corn starch. The molecular basis of artificial and spontaneous waxy mutants has also been clarified (Wessler and Varagona 1985). Several mutants arose by insertion of transposable elements into this gene. The molecular genetics of GBSS I has also been studied in rice (Hirano and Sano 1991; Hirano et al. 1998; Isshiki et al. 1998; Olsen and Purugganan 2002), barley (Domon et al. 2002a, b), and sorghum (McIntyre et al. 2008; Sattler et al. 2009). In rice, two dominant waxy alleles, $Wx^a$ and $Wx^b$ were observed mainly in indica and japonica rice groups, respectively. It was reported that $Wx^b$ arose by a point

mutation of the 5′ end of intron 1, resulting in aberrant splicing of the intron (Hirano et al. 1998; Isshiki et al. 1998; Olsen and Purugganan 2002). It was also reported that waxy rice originated by a 23-bp duplication of exon 2 (Wanchana et al. 2003). In barley, deletion in the 5′ region of the gene was found in wx genotype (Domon et al. 2002a, b; Patron et al. 2002). In sorghum, two different mutations were found for wx alleles; one resulting from a transposable element being inserted into the gene and the other from a point mutation resulting in amino acid substitution (McIntyre et al. 2008; Sattler et al. 2009). In hexaploid wheat, waxy wheat was artificially synthesized (Nakamura 1995), and it was concluded that deletion of the gene was responsible for the phenotype (Vrinten and Nakamura 1999).

Waxy phenotypes have also been observed in millets. In foxtail millet, the molecular basis of naturally occurring *wx* mutants has been well characterized (Nakayama et al. 1998; Fukunaga et al. 2002a; Kawase et al. 2005; Van et al. 2008). Waxy foxtail millet probably evolved from the non-waxy type after domestication, since the wild ancestor has a non-waxy endosperm (Nakayama et al. 1998). In addition to those two types, an intermediate or low-amylose type of this crop has been reported (Sakamoto 1987). Amylose content is positively correlated with amounts of GBSS 1 protein among the three phenotypes (Afzal et al. 1996) and is genetically controlled by waxy (GBSS 1) alleles (Nakayama et al. 1998). No other genes that regulate amylose content, such as *du* in rice (Okuno et al. 1983), are known in *S. italica*. The sequence of the full-length cDNA and the genomic structure of the *waxy* (*GBSS 1)* gene in foxtail millet was determined, and a preliminary diversity analysis indicated multiple origins of the waxy endosperm types (Fukunaga et al. 2002a). Kawase et al. (2005) analyzed 841 landraces of foxtail millet and classified 11 types by PCR-based methods. They concluded that waxy foxtail millet originated four times independently and low-amylose foxtail millet three times by insertions of transposable elements (Fig. 7.2). Van et al. (2008) found that the foxtail millet *waxy* gene has several SNPs and small indels. Recently, Hachiken et al. (2013) also investigated sequence variation of

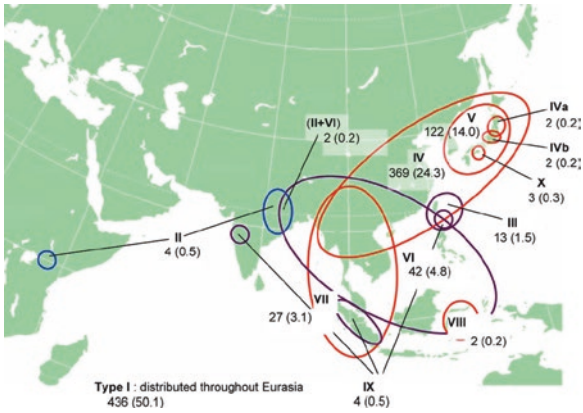


**Fig. 7.2** Summary of the geographical distribution of waxy and low amylose types of foxtail millet in Asia, Europe, and Africa (Kawase et al. 2005)

the *waxy* gene and revealed that *waxy* alleles of non-waxy accessions are more polymorphic than those of waxy and low-amylose accessions at the sequence level. This supports the hypothesis that waxy and low-amylose types originated from a non-waxy type.

### 7.4.2 *Variation in Phenol Color Reaction (Phr) and Evolution of the* Polyphenol Oxidase *(*PPO*) Gene*

Phenol color reaction (Phr) is a coloration of hulls/lemma and palea (grains) of cereals after soaking in phenol solution, as reported for rice (Oka 1953) and barley (Takeda and Chang 1996). The positive Phr type shows a black coloration after soaking in phenol solution, whereas the negative Phr type does not show coloration. Variation of Phr and geographical distribution of Phr phenotypes for foxtail millet have been reported (Kawase and Sakamoto 1982). It was shown in that study that Phr in foxtail millet is controlled by a single gene (positive Phr being dominant and negative Phr being recessive) and that the negative Phr type is predominant in Eurasia, whereas the positive Phr type generally has a skewed distribution toward subtropical and tropical regions including Nansei Islands of Japan, Taiwan, the Philippines, Nepal, and India (21–100 %).

The molecular mechanism of this trait has been investigated recently (Inoue et al. 2015). The *polyphenol oxidase* (*PPO*) gene responsible for Phr was isolated and molecular genetic basis of negative Phr and crop evolution of foxtail millet was investigated. Firstly, they searched for *PPO* gene homologs in a foxtail millet genome database (Bennetzen et al. 2012) using a rice *PPO* gene as a query, and successfully found three copies of the *PPO* gene. One of the *PPO* gene homologs on chromosome 7 showed the highest similarity with *PPO* genes expressed in hulls (grains) of other cereal species including rice, wheat, and barley and was designated as *Si7PPO*. Phr phenotypes and *Si7PPO* genotypes completely co-segregated in a segregating population. They also analyzed the genetic variation conferring negative Phr reaction. Of 480 accessions of the landraces investigated, 87 (18.1 %) showed positive Phr and 393 (81.9 %) showed negative Phr. In the 393 Phr negative accessions, three types of loss-of-function *Si7PPO* genes were predominant, with mutations found in various locations. One of them has an SNP in exon 1 resulting in a premature stop codon and was designated as stop codon type, another has an insertion of a transposon (*Si7PPO-TE1*) in intron 2 and was designated as TE1-insertion type, and the other has a 6-bp duplication in exon 3 resulting in the duplication of two amino acids and was designated as 6-bp duplication type. As a rare variant of the stop codon type, one accession additionally has an insertion of a transposon, *Si7PPO-TE2*, in intron 2 and was designated as "stop codon + TE2 insertion type." The geographical distribution of accessions with positive Phr and those with three major types of negative Phr was also investigated (Fig. 7.3). Accessions with positive Phr were found in subtropical

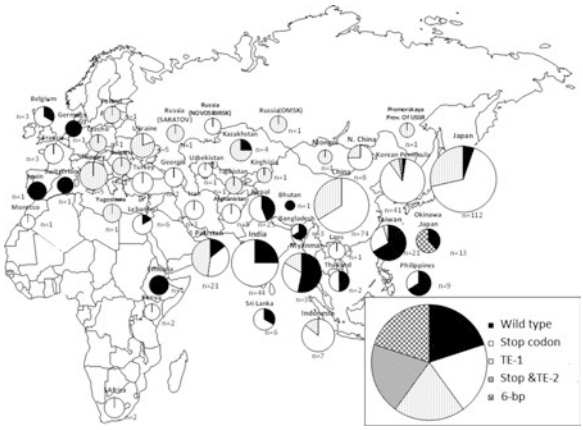


**Fig. 7.3** Geographical distribution of positive Phr (wild type) and four different genotypes of Phr, stop codon type, TE1 insertion type, 6-bp duplication type, and stop codon and TE2 insertion type (Inoue et al. 2015)

and tropical regions at frequencies of ca. 25–67 % and those with negative Phr were broadly found in Europe and Asia. The stop codon type was found in 285 accessions and was broadly distributed in Europe and Asia, whereas the TE1-insertion type was found in 99 accessions from Europe and Asia but was not found in India. The 6-bp duplication type was found in only eight accessions from Nansei Islands (Okinawa Prefecture) of Japan. They also analyzed Phr in the wild ancestor and concluded that the negative Phr type was likely to have originated after domestication of foxtail millet. Their study also suggested that the negative Phr of foxtail millet arose by multiple independent loss of function of *PPO* gene, that proved advantageous under some environmental conditions and under human selection, as also seen in rice (Yu et al. 2008) and barley (Taketa et al. 2010).

## 7.5 Perspective

Recent studies in phylogeny and association mapping using Next-Generation Sequencing (NGS) technology (Jia et al. 2013) have updated relationships in foxtail millet and revealed several candidate genes involved in domestication and diversification of landraces in foxtail millet. Genetic mapping of a gene involved in panicle morphology (Sato et al. 2013) and QTLs for inflorescence structure, branching and height (Doust et al. 2004; Doust et al. 2005; Mauro-Herrera and Doust 2016), and flowering time (Mauro-Herrera et al. 2013) also have been carried out. Some of these studies on association mapping and QTLs are described in chapter 12. Further analyses of candidate genes in the context of domestication and crop evolution of foxtail millet by using foxtail millet landraces from broad area of Europe and Asia will be interesting. Currently, we are focused on genes involved in panicle morphology, which will be helpful to understand crop evolution and diversification of foxtail millet.

## References

Afzal M, Kawase M, Nakayama H, Okuno K. Variation in electrophoregrams of total seed protein and Wx protein in foxtail millet. In: Janick J, editor. Progress in new crops. Alexandria, VA: ASHS Press; 1996. p. 191–5.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30:555–61.

Collins GN. A new type of Indian corn from China. USDA Bur Pl Ind Bull. 1909;16:1–30.

de Wet JMJ, Oestry-Stidd LL, Cubero JI. Origins and evolution of foxtail millet (*Setaria italica*). Journ d'Agric et de Bot. 1979;26:53–64.

Dekaprelevich LL, Kasparian AS. A contribution to the study of foxtail millet (*Setaria italica* P.B. *maxima*. Alef.) cultivated in Georgia (western Transcaucasia). Bull Appl Bot Plant Breed. 1928;19:533–72.

Doebley J, Stec A, Hubbard L. The evolution of apical dominance in maize. Nature. 1997;386:485–8.

Domon E, Fujita M, Ishikawa N. The insertion/deletion polymorphisms in the *Waxy* gene of barley genetic resources from East Asia. Theor Appl Genet. 2002a;104:132–8.

Domon E, Saito A, Takeda K. Comparison of the waxy locus sequence from a non-waxy strain and two waxy mutants of spontaneous and artificial origins in barley. Genes Genet Syst. 2002b;77 :351–9.

Doust AD, Devos KM, Gadberry MD, Gale MD, Kellogg EA. Genetic control of branching in foxtail millet. Proc Natl Acad Sci U S A. 2004;101:9045–50.

Doust AD, Devos KM, Gadberry MD, Gale MD, Kellogg EA. The genetic basis for inflorescence variation between foxtail and green millet (Poaceae). Genetics. 2005;169:1659–72.

Eda M, Izumitani A, Ichitani K, Kawase M, Fukunaga K. Geographical variation of foxtail millet, *Setaria italica* (L.) P. Beauv. based on rDNA PCR–RFLP. Genet Res Crop Evol. 2013;60:265–74.

Fogg WH. Swidden cultivation of foxtail millet by Taiwan aborigines: a cultural analogue of the domestication of *Setaria italica* in China. In: Keighty DN, editor. The origins of Chinese civilization. Berkeley: University of California Press; 1983. p. 95–115.

Fukunaga K, Kato K. Mitochondrial DNA variation in foxtail millet, *Setaria italica* (L.) P. Beauv. Euphytica. 2003;129:7–13.

Fukunaga K, Domon E, Kawase M. Ribosomal DNA variation in foxtail millet, Setaria italica (L.) P. Beauv. and a survey of variation from Europe and Asia. Theor Appl Genet. 1997a;97:751–6.

Fukunaga K, Kawase M, Sakamoto S. Variation of caryopsis length and width among landraces of foxtail millet, *Setaria italica* (L.) P. Beauv. Jpn J Trop Agric. 1997b;41:235–40.

Fukunaga K, Kawase M, Kato K. Structural variation in the *Waxy* gene and differentiation in foxtail millet [*Setaria italica* (L.) P. Beauv.]: implications for multiple origins of the waxy phenotype. Mol Genet Genomics. 2002a;268:214–22.

Fukunaga K, Wang ZM, Kato K, Kawase M. Geographical variation of nuclear genome RFLPs and genetic differentiation in foxtail millet, *Setaria italica* (L.) P. Beauv. Genet Res Crop Evol. 2002b;49:95–101.

Fukunaga K, Ichitani K, Taura S, Sato M, Kawase M. Ribosomal DNA intergenic spacer sequence in foxtail millet, *Setaria italica* (L.) P. Beauv. and its characterization and application to typing of foxtail millet landraces. Hereditas. 2005;142:38–44.

Fukunaga K, Ichiatani K, Kawase M. Phylogenetic analysis of rDNA intergenic spacer subrepeats and its implication for domestication history of foxtail millet, *Setaria italica*. Theor Appl Genet. 2006;113:261–9.

Fukunaga K, Ichitani K, Kawase M. rDNA polymorphism of foxtail millet (*Setaria italica* ssp. *italica*) landraces in northern Pakistan and Afghanistan and in its wild ancestor (*S. italica* ssp. *viridis*). Genet Resour Crop Evol. 2011;58:825–30.

Hachiken T, Sato K, Hasegawa T, Ichitani K, Kawase M, Fukunaga K. Geographic distribution of *Waxy* gene SNPs and indels in foxtail millet, *Setaria italica* (L.) P. Beauv. Genet Res Crop Evol. 2013;60:1559–70.

Hammer K, Khoshbakht K. Foxtail millet (*Setaria italica* (L.) P. Beauv.) in Mazandaran/Northern Iran. Genet Res Crop Evol. 2007;54:907–11.

Harlan JR. Crops and man. Madison, WI: American Society of Agronomy, Crop Science Society of America; 1975.

Hirano HY, Sano Y. Molecular characterization of the waxy locus of rice (*Oryza sativa*). Plant Cell Physiol. 1991;32:989–97.

Hirano HY, Eighuchi M, Sano Y. A single base change altered the regulation of the *Waxy* gene at the posttranscriptional level during the domestication of rice. Mol Biol Evol. 1998;15:978–87.

Hirano R, Naito K, Fukunaga K, Watanabe KN, Ohsawa R, Kawase M. Genetic structure of landraces in foxtail millet (*Setaria italica* (L.) P. Beauv.) revealed with transposon display and interpretation to crop evolution of foxtail millet. Genome. 2011;54:498–506.

Hunt HV, Linden MV, Liu X, Motuzaite-Matuzeviciute G, Colledge S, Jones MK. Millets across Eurasia: chronology and context of early records of the genera *Panicum* and *Setaria* from archaeological sites in the Old World. Veg Hist Archaeobot. 2008;17(Suppl):5–18.

Inoue T, Yuo T, Ohta T, Hitomi E, Ichitani K, Kawase M, et al. Multiple origins of the phenol reaction negative phenotype in foxtail millet, *Setaria italica* (L.) P. Beauv, were caused by independent loss-of-function mutations of the *polyphenol oxidase* (*Si7PPO*) gene during domestication. Mol Genet Genomics. 2015;290:1563–74.

Ishii T, Numaguchi K, Miura K, Yoshida K, Thanh PT, Htun TM, et al. *OsLG1* regulates a closed panicle trait in domesticated rice. Nat Genet. 2013;45:462–5.

Isshiki M, Morino K, Nakajima M, Okagaki RJ, Wessler SR, Izawa T, et al. A naturally occurring functional allele of the rice waxy locus has a GT to TT mutation at the 5′ splice site of the first intron. Plant J. 1998;15:133–8.

Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013;45:957–61.

Jusuf M, Pernes J. Genetic variability of foxtail millet (*Setaria italica* P. Beauv.). Theor Appl Genet. 1985;71:385–93.

Kato S, Kosaka H, Hara S. On the affinity of rice varieties shown by the fertility of hybrid plants. Rep Bul Sci Fak Terkult Kyushu Imp Univ. 1928;3:132–47 (in Japanese with English summary).

Kawase M. Genetic variation and landrace differentiation of foxtail millet, *Setaria italica*, in Eurasia. Ph.D. thesis. Kyoto University, Japan, 1986.

Kawase M, Fukunaga K. Distribution of Type D, a landrace group newly determined by means of hybrid sterility in foxtail millet, *Setaria italica* (L.) P. Beauv. Breed Res. 1999;1:302 (in Japanese).

Kawase M, Sakamoto S. Geographical distribution and genetic analysis of phenol color reaction in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1982;63:117–9.

Kawase M, Sakamoto S. Variation, geographical distribution and genetical analysis of esterase isozymes in foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1984;67:529–33.

Kawase M, Sakamoto S. Geographical distribution of landrace groups classified by hybrid pollen sterility in foxtail millet, *Setaria italica* (L.) P. Beauv. J Jpn Breed. 1987;37:1–9.

Kawase M, Ochiai Y, Fukunaga K. Characterization of foxtail millet, *Setaria italica* (L.) P. Beauv. in Pakistan based on intraspecific hybrid pollen sterility. Breed Sci. 1997;47:45–9.

Kawase M, Fukunaga K, Kato K. Diverse origins of waxy foxtail millet crops in East and Southeast Asia mediated by multiple transposable element insertions. Mol Genet Genomics. 2005;274:131–40.

Kihara H, Kishimoto E. Bastarde zwischen *Setaria italica* und *S. viridis*. Bot Mag. 1942;20:63–7 (in Japanese with German summary).

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, et al. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci U S A. 2007;104:1424–9.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, et al. An SNP caused loss of seed shattering during rice domestication. Science. 2006;312:1392–6.

Kűster H. Neolithic plant remains from Eberdingenhochdo, southern Germany. In: Van Zeist WV, Casparoe WA, editors. Plants and ancient man (studies in palaeoethnobotany). Rotterdam: AA Bakame; 1984. p. 307–11.

Le Thierry d'Ennequin M, Panaud O, Toupance B, Sarr A. Assessment of genetic relationships between *Setaria italica* and its wild relative *S. viridis* using AFLP markers. Theor Appl Genet. 2000;100:1061–6.

Li Y, Wu SZ. Traditional maintenance and multiplication of foxtail millet (*Setaria italica* (L.) P. Beauv) landraces in China. Euphytica. 1996;87:33–8.

Li HW, Li CH, Pao WK. Cytological and genetical studies of the interspecific cross of the cultivated foxtail millet, *Setaria italica* (L.) Beauv., and the green foxtail millet, *S. viridis*. J Am Soc Agron. 1945;37:32–54.

Li Y, Cao YS, Wu SZ, Zhang XZ. A diversity analysis of foxtail millet (*Setaria italica* (L.) P. Beauv.) landraces of Chinese origin. Genet Resour Crop Evol. 1995a;45:279–85.

Li Y, Wu SZ, Cao YS. Cluster analysis of an international collection of foxtail millet (*Setaria italica* (L.) P. Beauv). Euphytica. 1995b;83:79–85.

Li Y, Jia J, Wang W, Wu S. Intraspecific and interspecific variation in *Setaria* revealed by RAPD analysis. Genet Resour Crop Evol. 1998;45:279–85.

Lisitsina A. Main types of ancient farming on the Caucasus on the basis of palaeo-ethnobotanical research. Ber Deut Bot Ges. 1976;91:47–57.

Mauro-Herrera M, Doust AN. Development and genetic control of plant architecture and biomass in the panicoid grass, *Setaria*. PLoS One. 2016;11(3):e0151346.

Mauro-Herrera M, Wang X, Barbier H, Brutnell TP, Devos KM, Doust AN. Genetic control and comparative genomic analysis of flowering time in *Setaria* (Poaceae). Genes, Genomes, and Genetics. 2013;3:283–95.

McIntyre CL, Drenth J, Gonzalez N, Henzell RG, Jordan DR. Molecular characterization of the waxy locus in sorghum. Genome. 2008;51:524–33.

Nakamura T. Production of waxy (amylose-free) wheats. Mol Gen Genet. 1995;248:253–9.

Nakayama H, Afzal M, Okuno K. Intraspecific differentiation and geographical distribution of Wx alleles for low amylase content in endosperm of foxtail millet, *Setaria italica* (L.) Beauv. Euphytica. 1998;102:289–93.

Nakayama H, Namai H, Okuno K. Geographical variation of the alleles at the two prolamin loci, *Pro1* and *Pro2*, in foxtail millet, *Setaria italica* (L.) P. Beauv. Genes Genet Syst. 1999;74:293–7.

Nasu H, Momohra A, Yasuda Y, He J. The occurrence and identification of *Setaria italica* (L.) P. Beauv. (foxtail millet) grains from the Chengtoushan site (ca. 5800 cal B.P.) in central China, with reference to the domestication centre in Asia. Veg Hist Archaeobot. 2007;16:481–94.

Nguyen Van F, Pernes J. Genetic diversity of foxtail millet (*Setaria italica*). In: Jacquard P, editor. Genetic differentiation and dispersal in plants. NATO ASI Series, vol. G5. Berlin: Springer; 1985. p. 113–128.

Ochiai Y. Variation in tillering and geographical distribution of foxtail millet (*Setaria italica* P. Beauv.). Breed Sci. 1996;46:143–6.

Ochiai Y, Kawase M, Sakamoto S. Variation and distribution of foxtail millet (*Setaria italica* P. Beauv.) in the mountainous areas of northern Pakistan. Breed Sci. 1994;44:413–8.

Oka HI. Phylogenetic differentiation of the cultivated rice plant. 1. Variations in respective characteristics and their combinations in rice cultivars. Jpn J Breed. 1953;3:33–43 (in Japanese with English summary).

Okuno K, Fuwa H, Yano M. A new mutant gene lowering amylose content in endosperm starch in rice, *Oryza sativa* L. Japan J Breed. 1983;33:387–94.

Olsen KM, Purugganan MD. Molecular evidence on the origin and evolution of glutinous rice. Genetics. 2002;162:941–50.

Patron NJ, Smith AM, Fahy BF, Hylton CM, Naldrett MJ, Rossnagel BG, et al. The altered pattern of amylose accumulation in the endosperm of low-amylose barley cultivars is attributable to a

single mutant allele of granule-bound starch synthase I with a deletion in the 5-non-coding region. Plant Physiol. 2002;130:190–8.
Prasada Rao KE, de Wet JMJ, Brink DE, Mengesha MH. Infraspecific variation and systematics of cultivated *Setaria italica*, foxtail millet (Poaceae). Econ Bot. 1987;41:108–16.
Sakamoto S. Origin and dispersal of common millet and foxtail millet. Japan Agr Res Quart. 1987;21:84–9.
Sakamoto S. Glutinous-endosperm starch food culture specific to Eastern and Southeastern Asia. In: Ellen R, Fukui K, editors. Redefining nature: ecology, culture and domestication. Oxford: Berg; 1996. p. 215–31.
Sano Y. Differential regulation of gene expression in rice endosperm. Theor Appl Genet. 1984;68:467–73.
Sato K, Mukainari Y, Naito K, Fukunaga K. Construction of a foxtail millet linkage map and mapping *spikelet-tipped bristles 1*(*stb1*) by using transposon display markers and simple sequence repeat markers with genome sequence information. Mol Breed. 2013;31:675–84.
Sattler SE, Singh J, Haas EJ, Guo L, Sarath G, Pedersen JF. Two distinct waxy alleles impact the granule-bound starch synthase in sorghum. Mol Breed. 2009;24:349–59.
Schontz D, Rether B. Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv. RFLP using a heterologous rDNA probe. Plant Breed. 1998;117:231–4.
Schontz D, Rether B. Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv.: Identification and classification of lines with RAPD markers. Plant Breed. 1999;118:190–2.
Takeda K, Chang CL. Inheritance and geographical distribution of phenol reaction-less varieties of barley. Euphytica. 1996;90:217–21.
Takei E. Characteristics and ethnobotany of millets in the Southwestern (Nansei) Islands of Japan (in Japanese with English abstract). Ph.D. thesis, Faculty of Agriculture, Kyoto University, 1994.
Takei E, Sakamoto S. Geographical variation of heading response to daylength in foxtail millet (*Setaria italica* P. Beauv.). Jpn J Breed. 1987;37:150–8.
Takei E, Sakamoto S. Further analysis of geographical variation of heading response to daylength in foxtail millet (*Setaria italica* P. Beauv.). Japan J Breed. 1989;39:285–98.
Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, Kakeda K, et al. Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. Proc Natl Acad Sci U S A. 2008;105:4062–7.
Taketa S, Matsuki K, Amano S, Saisho D, Himi E, Shitsukawa N, et al. Duplicate polyphenol oxidase genes on barley chromosome 2H and their functional differentiation in the phenol reaction of spikes and grains. J Exp Bot. 2010;61:3983–93.
Van S, Onoda S, Kim MY, Kim KD, Lee SH. Allelic variation of the Waxy gene in foxtail millet [*Setaria italica* (L.) P. Beauv.] by single nucleotide polymorphisms. Mol Genet Genomics. 2008;279:255–66.
Vavilov NI. Studies on the origin of cultivated plants. Inst Appl Bot Plant Breed. 1926;16:1–248.
Vrinten P, Nakamura T. Molecular characterization of waxy mutations in wheat. Mol Gen Genet. 1999;261:463–71.
Wanchana S, Toojinda T, Tragoonrung S, Vanavichit A. Duplicated coding sequence in the waxy allele of tropical glutinous rice (*Oryza sativa* L.). Plant Sci. 2003;165:1193–9.
Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, et al. The origin of the naked grains of maize. Nature. 2005;436:714–9.
Wessler SR, Varagona MJ. Molecular basis of mutations at Waxy locus of maize: correlation with the fine structure genetic map. Proc Natl Acad Sci U S A. 1985;82:4177–81.
Yoshida S. Wild food plants and vegeculture. In: Yoshida S, Matthews PJ, editors. Vegeculture in Eastern Asia and Oceania, JCAS Symposium Series, vol. 16. Osaka: National Museum of Ethnology; 2002. p. 31–44.
Yu Y, Tang T, Qian Q, Wang Y, Yan M, Zeng D, et al. Independent losses of function in a polyphenol oxidase in rice: differentiation in grain discoloration between subspecies and the role of positive selection under domestication. Plant Cell. 2008;20:2946–59.

# Part II
# Genomics

# Chapter 8
# Genome Structure and Comparative Genomics

**Katrien M. Devos, Xiaomei Wu, and Peng Qi**

**Abstract** Two genome assemblies have been generated for *Setaria italica*, one in the accession Yugu1 and one in the accession Zhang gu. A comparison of the two assemblies showed overall high levels of colinearity, but identified a number of likely misoriented and/or misplaced sequence scaffolds. Using available comparative information between Setaria, rice, sorghum, *Brachypodium,* and switchgrass, we reconstructed the structure of the 12-chromosome ancestral grass genome from which the grasses radiated some 70 million years ago. Comparing the structure of the ancestral grass genome with that of current-day Setaria and sorghum showed that the Setaria genome has undergone more rearrangements than sorghum. The reasons for the greater stability of the sorghum compared to the Setaria genome are unknown.

**Keywords** *Setaria italica* • Foxtail millet • Ancestral grass genome • Sorghum • Comparative genomics • Genome organization

## 8.1 Introduction

Foxtail millet, *Setaria italica*, is the second most important small millet after pearl millet. It is grown for the production of grain as well as forage. As such, it has received considerable research attention, particularly in China, where foxtail millet is predominantly cultivated. In recent years, foxtail millet together with its wild progenitor, *S. viridis* (green foxtail), have also gained attention as model systems for more genetically complex relatives such as the biofuel crop switchgrass (*Panicum*

K.M. Devos (✉) • P. Qi
Institute of Plant Breeding, Genetics and Genomics (Dept. of Crop and Soil Sciences), and Dept. of Plant Biology, University of Georgia, Athens, GA 30602, USA
e-mail: kdevos@uga.edu

X. Wu
Institute of Plant Breeding, Genetics and Genomics (Dept. of Crop and Soil Sciences), and Dept. of Plant Biology, University of Georgia, Athens, GA 30602, USA

Center for Plant Environmental Sensing, and College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 310036, China



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_8

*virgatum*). *S. italica* and *S. viridis* are inbreeding diploid species ($2n=2x=18$) with a relatively small genome (1C = ~515 Mb), and both species have been sequenced. Genome assemblies anchored to genetic maps are available for two *S. italica* cultivars, Yugu1 (Bennetzen et al. 2012) and Zhang gu (Zhang et al. 2012) and resequencing data is available for several hundred *S. italica* (Jia et al. 2013) and *S. viridis* accessions (genome.jgi.doe.gov/DevoffC4grasses).

A species' value as a model is enhanced significantly if the structural relationship between the genomes of the model and target crops is known. Structural comparative genomics in plants has its roots in the colinearity studies of the late 1980s and 90s that were based on common markers on genetics maps (Gale and Devos 1998; Bennetzen and Freeling 1997; Choi et al. 2004). The first map-based comparative analysis involving foxtail millet was published in 1998 and established the relationship between the Setaria and rice genomes (Devos et al. 1998). Despite the fact that rice and foxtail millet had different chromosome numbers (12 and 9, respectively), the genomes of both species showed a high level of colinearity. The sequencing of the Setaria genome paved the way for much more detailed comparative analyses (Bennetzen et al. 2012; Zhang et al. 2012). Here, in addition to describing the structural relationship between the genomes of foxtail millet, rice, and sorghum, we used comparative data between these three species, switchgrass (Daverdin et al. 2015) and *Brachypodium* to reconstruct the ancestral grass genome, and trace the evolutionary history from ancestral to extant foxtail millet chromosomes.

## 8.2 The Yugu1 and Zhang gu Foxtail Millet Genome Assemblies

The Yugu1 genome assembly was obtained from an 8.3X shotgun sequence generated using Sanger sequencing technology (Bennetzen et al. 2012). The main genome assembly comprised 405.7 Mb organized in 336 scaffolds with 98.9 % of the sequence incorporated into nine scaffolds (nine pseudomolecules). Overall, the contig and scaffold N50 (L50) were 982 (126.3 kb) and 4 (47.3 Mb), respectively. Scaffolds were anchored to a genetic map comprising 992 single nucleotide polymorphism (SNP) markers. Because linkage groups had previously been mapped to the foxtail millet karyotype (Wang et al. 1998), this information was used to assign the nine pseudomolecules to specific foxtail millet chromosomes labeled in roman numericals, I to IX.

The Zhang gu genome was shotgun sequenced to a depth of 80X on Illumina Genome Analyzer II and HiSeq 2000 platforms. The assembly totaled 423 Mb and had a contig and scaffold N50 (L50) of 4667 (25 kb) and 136 (1 Mb), respectively. The 613 longest scaffolds, representing approximately 95 % of the sequence, were anchored to a 751 loci genetic map.

In order to compare the structure of the two Setaria genomes, the Yugu1 assembly v2.0 was downloaded from Phytozome (phytozome.jgi.doe.gov) and the Zhang gu chromosome assembly foxtail_millet.chr.fa.gz was downloaded

from http://foxtailmillet.genomics.org.cn. A dot plot comparison of the Yugu1 and Zhang gu assemblies showed overall high levels of colinearity. Nevertheless, the two assemblies varied by a number of rearrangements, in particular in centromeric and pericentromeric regions where recombination is low and genetic mapping data was of limited help in ordering and orienting scaffolds. Discrepancies between the Yugu1 and Zhang gu assemblies were limited to four or less regions in chromosomes I to V, but were more extensive in chromosomes VI to IX (Fig. 8.1). Regions larger than 300 kb that were rearranged between the Yugu1 and Zhang gu assemblies and their approximate coordinates in both sequence assemblies are listed in Table 8.1.

We looked in more detail at 22 inversions larger than 400 kb that differentiated the Yugu1 and Zhang gu sequences. Inversion 9_1 (Table 8.1; Fig. 8.1) had previously been identified as a Yugu1-specific inversion (Bennetzen et al. 2012) and was not further analyzed here. To pinpoint putative assembly errors, we identified the Yugu1 and Zhang gu scaffolds that were located in each of the inversions. Zhang gu scaffolds were retrieved from the "Millet_scaffoldVersion2.3.fa.gz" file downloaded from http://foxtailmillet.genomics.org.cn. Because the size of scaffolds retrieved from the foxtail millet database (http://foxtailmillet.genomics.org.cn) through Mapview differed from the size of scaffolds with the same ID in the downloaded "Millet_scaffoldVersion2.3.fa.gz" file for 29 % of the scaffolds checked, we used an in-house script to count the number of bases in relevant scaffolds in the Millet_scaffoldVersion2.3.fa.gz file. If the size of an inversion corresponded, within a range of 20 kb, to the total size of all scaffolds that were located within the inverted region, we assumed that the inversion was caused by misorientation of those scaffolds. Based on this assumption, eight of the 21 inversions analyzed were likely caused by assembly errors in the Zhang gu sequence and one was caused by an incorrectly placed and oriented scaffold in the Yugu1 sequence (Table 8.1). Misplacement of this Yugu1 scaffold (block 5_3 in Table 8.1 and Fig. 8.1) was confirmed by a mapped SNP marker (K. M. Devos, unpublished data). The 20 kb range for identifying inversions that corresponded to complete scaffolds was a very strict criterion considering that the breakpoints listed in Table 8.1 denoted the minimum size of the inversions. Indeed, scrutiny of 12 inversions where scaffolds extended for more than 20 kb beyond the breakpoints showed that, in ten cases, the end points of the Zhang gu scaffolds were located within the breakpoint range (data not shown). One of the 12 inversions (block 8_5 in Table 8.1 and Fig. 8.1) was likely due to an incorrectly oriented scaffold in Yugu1. This scaffold extended for ~280 kb beyond the inversion. This ~280 kb region was colinear between Zhang gu and Yugu1, but comprised a single scaffold in Zhang gu. We surmised that 8_5 was misoriented in Yugu1 and that the neighboring ~280 kb region was misoriented in Zhang gu. The cause of inversion 3_1 could not be traced. Our results showed that rearranged segments typically corresponded to complete scaffolds indicating that the problem laid with the ordering and orienting of scaffolds and not with assemblies within scaffolds. Overall, out of the 21 putative misassemblies analyzed, 19 (90 %) likely occurred in the Zhang gu sequence and two (10 %) in Yugu1.

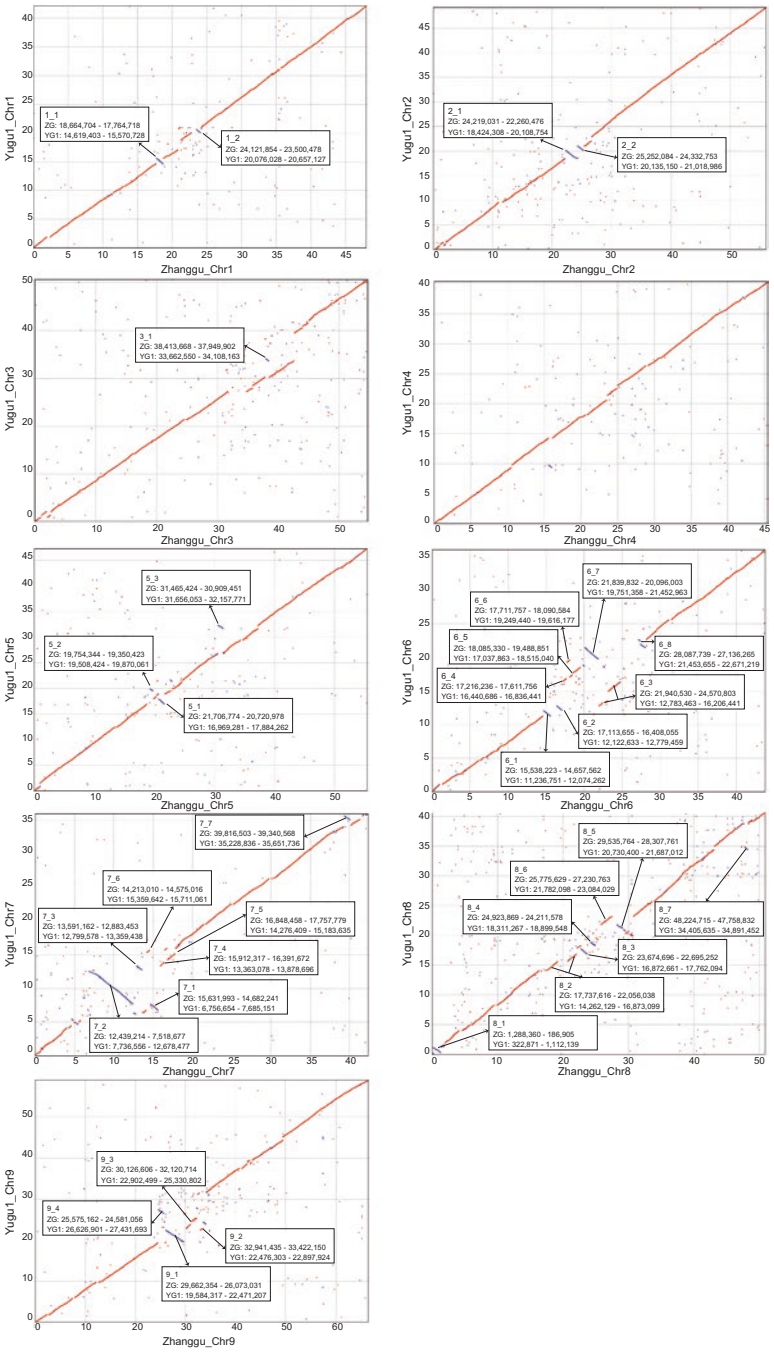


**Fig. 8.1** Dot plots showing the relationship between the Yugu1 and Zhang gu chromosome assemblies. Regions larger than 400 kb and their breakpoints are indicated

**Table 8.1** Regions >300 kb that are differentially placed and/or oriented in the Yugu1 compared to the Zhang gu chromosome assemblies

| Region ID | | Assembly error[a] | Yugu1 | | | Zhang gu | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Start | End | Length | Start | End | Length | Scaffold index | Scaffold length |
| 1_1 | – | *Zhang gu* | 14,619,403 | 15,570,728 | 951,325 | 18,664,704 | 17,764,718 | 899,986 | 105 | 940,479 |
| 1_2 | – | Zhang gu | 20,076,028 | 20,657,127 | 581,099 | 24,121,854 | 23,500,478 | 621,376 | 117 | 622,654 |
| 2_1 | – | *Zhang gu* | 18,424,308 | 20,108,754 | 1,684,446 | 24,219,031 | 22,260,476 | 1,958,555 | 287 | 1,874,024 |
| 2_2 | – | Zhang gu | 20,135,150 | 21,018,986 | 883,836 | 25,252,084 | 24,332,753 | 919,331 | 102 | 931,090 |
| 3_1 | – | ND | 33,662,550 | 34,108,163 | 445,613 | 38,413,668 | 37,949,202 | 464,466 | 132 | 1,230,272 |
| 5_1 | – | *Zhang gu* | 16,969,281 | 17,884,262 | 914,981 | 21,706,774 | 20,720,978 | 985,796 | 92 | 884,689 |
| 5_2 | – | NA | 19,508,424 | 19,870,061 | 361,637 | 19,754,344 | 19,350,423 | 403,921 | 290 | 1,605,301 |
| 5_3 | – | Yugu1 | 31,656,053 | 32,157,771 | 501,718 | 31,465,424 | 30,909,451 | 555,973 | 2 | 3,433,248 |
| 6_1 | – | Zhang gu | 11,236,751 | 12,074,262 | 837,511 | 15,538,223 | 14,657,562 | 880,661 | 120 | 891,560 |
| 6_2 | – | Zhang gu | 12,122,633 | 12,779,459 | 656,826 | 17,113,655 | 16,408,055 | 705,600 | 384 | 707,442 |
| 6_3 | + | NA | 12,783,463 | 16,206,441 | 3,422,978 | 21,940,530 | 24,570,803 | 2,630,273 | 134, 63, 348 | 2,076,409 |
| 6_4 | + | NA | 16,440,686 | 16,836,441 | 395,755 | 17,216,236 | 17,611,756 | 395,520 | 333 | 398,101 |
| 6_5 | + | NA | 17,037,863 | 18,515,040 | 1,477,177 | 18,085,330 | 19,488,851 | 1,403,521 | 221(P), 291, 247 | |
| 6_6 | + | NA | 19,249,440 | 19,616,177 | 366,737 | 17,711,757 | 18,090,584 | 378,827 | 221 | 736,944 |
| 6_7 | – | Zhang gu | 19,751,358 | 21,452,963 | 1,701,605 | 21,839,832 | 20,096,003 | 1,743,829 | 202 | 1,750,473 |
| 6_8 | – | *Zhang gu* | 21,453,655 | 22,671,219 | 1,217,564 | 28,087,739 | 27,136,265 | 951,474 | 242 | 1,219,385 |
| 7_1 | – | *Zhang gu* | 6,756,654 | 7,685,151 | 928,497 | 15,631,993 | 14,682,241 | 949,752 | 113 | 1,037,179 |

(continued)

**Table 8.1** (continued)

| Region ID | | Assembly error[a] | Yugu1 | | | Zhang gu | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Start | End | Length | Start | End | Length | Scaffold index | Scaffold length |
| 7_2 | – | *Zhang gu* | 7,736,556 | 12,678,477 | 4,941,921 | 12,439,214 | 7,518,677 | 4,920,537 | 509, 262, 412, 171 | 4,570,300 |
| 7_3 | – | *Zhang gu* | 12,799,578 | 13,359,438 | 559,860 | 13,591,162 | 12,883,453 | 707,709 | 101 | 765,901 |
| 7_4 | + | NA | 13,363,078 | 13,878,696 | 515,618 | 15,912,317 | 16,391,672 | 479,355 | 268 | 481,785 |
| 7_5 | + | NA | 14,276,409 | 15,183,635 | 907,226 | 16,848,458 | 17,757,779 | 909,321 | 327 | 909,795 |
| 7_6 | + | NA | 15,359,642 | 15,711,061 | 351,419 | 14,213,010 | 14,575,016 | 362,006 | 572 | 362,495 |
| 7_7 | – | *Zhang gu* | 35,228,836 | 35,651,736 | 422,900 | 39,816,503 | 39,340,568 | 475,935 | 276 | 621,760 |
| 8_1 | – | Zhang gu | 322,871 | 1,112,139 | 789,268 | 1,288,360 | 186,905 | 1,101,455 | 751 | 1,099,786 |
| 8_2 | + | NA | 14,262,129 | 16,873,099 | 2,610,970 | 17,737,616 | 22,056,038 | 4,318,422 | 97, 318, 397, 66 | 2,577,865 |
| 8_3 | – | *Zhang gu* | 16,872,661 | 17,762,094 | 889,433 | 23,674,696 | 22,695,252 | 979,444 | 4 | 1,579,811 |
| 8_4 | – | Zhang gu | 18,311,267 | 18,899,548 | 588,281 | 24,923,869 | 24,211,578 | 712,291 | 52 | 731,020 |
| 8_5 | – | *Yugu1* | 20,730,400 | 21,687,012 | 485,817 | 29,535,764 | 28,307,761 | 1,228,003 | 158, 182 | 1,645,605 |
| 8_6 | + | NA | 21,782,098 | 23,084,029 | 1,301,931 | 25,775,629 | 27,230,763 | 1,455,134 | 107 | 1,733,986 |
| 8_7 | – | Zhang gu | 34,405,635 | 34,891,452 | 485,817 | 48,224,715 | 47,758,832 | 465,883 | 688 | 470,489 |
| 9_1 | – | Yugu1[b] | 19,584,317 | 22,471,207 | 2,886,890 | 29,662,354 | 26,073,031 | 3,589,323 | 98, 356 | 1,090,704 |
| 9_2 | + | NA | 22,476,303 | 22,897,924 | 421,621 | 32,941,435 | 33,422,150 | 480,715 | 443 | 552,083 |
| 9_3 | + | NA | 22,902,499 | 25,330,802 | 2,428,303 | 30,126,606 | 32,120,714 | 1,994,108 | 353, 95 | 1,813,740 |
| 9_4 | – | *Zhang gu* | 26,626,901 | 27,431,693 | 804,792 | 25,575,162 | 24,581,056 | 994,106 | 236 | 249,103 |

[a]NA (not analyzed): only inversions larger than 400 kb were analyzed; ND (not determined): inversions did not correspond to complete scaffolds and hence the source of the inversion could not be traced; Accession names in italics indicate that the scaffold size in that accession extended for more than 20 kb beyond the inversion breakpoints listed (which represented the minimum size of the inversion) but scaffold ends were nevertheless located within breakpoint ranges

[b]Yugu1-specific inversion (Bennetzen et al. 2012)

## 8.3 Genome Organization

The Setaria genome has the organizational structure typically seen in small plant genomes with the majority of the repeats, in particular the gypsy retrotransposons, being located in the low recombinant centromeric and pericentromeric regions. Approximately 40–46 % of the DNA has been classified as repetitive (Bennetzen et al. 2012; Zhang et al. 2012, chapter 9), a number that is in line with that found in other plant species with comparable genome sizes. Genes are enriched in the distal regions of the chromosome arms. Estimations for the number of protein-coding genes in the Setaria genome vary from 34,584 in the Yugu1 sequence (Bennetzen et al. 2012) to 38,801 in the Zhang gu sequence (Zhang et al. 2012). Manual validation of a subset of these genes, however, indicated that the true gene number may be closer to 29,000 (Bennetzen et al. 2012).

## 8.4 Relationships Between the Genomes of Foxtail Millet, Sorghum, Rice, and Brachypodium

The first comparative structural analysis between the genomes of foxtail millet ($n=9$) and rice ($n=12$) was conducted in 1998 using a set of common markers that had been genetically mapped in both species (Devos et al. 1998). Although the number of comparative data points was limited to 87, this study nevertheless demonstrated that foxtail millet chromosomes I, IV, V, VI, and VIII corresponded to rice chromosomes 2, 6, 1, 8, and 11, respectively, while foxtail millet chromosomes II, III, VII, and IX each corresponded to two complete or partial rice chromosomes (foxtail II: rice 7 and 9; foxtail III: rice 5 and most of rice 12; foxtail VII: rice 4 and the distal region of the short arm of rice 12; and foxtail IX: rice 3 and 10). One-to-two relationships were generally formed by the insertion of one rice chromosome into the pericentromeric region of another chromosome. The exception was foxtail millet chromosome VII, where the rice segment which corresponded to the region of rice chromosome 12 that was duplicated on rice chromosome 11 (Nagamura et al. 1995) was fused to the telomere of the long arm of rice chromosome 4.

Following the completion of the genome sequences of foxtail millet, more detailed comparative analyses were carried out. While the genomic comparisons more precisely defined the previously identified synteny breakpoints, very few additional rearrangements were found (Bennetzen et al. 2012). Newly identified syntenic blocks included the distal ~2 Mb at the top of Setaria chromosome III which was syntenic to the distal region of the long arm of rice chromosome 4, and an ~1.5 Mb region, located between the regions with synteny to rice chromosomes 4 and 12 at the bottom of Setaria chromosome VII, which was syntenic to an interstitial region on rice chromosome 5 (Fig. 8.2 and Table 8.2). Setaria chromosomes III and VII thus each had synteny to segments of the same three rice chromosomes. In addition, at least 34 inversions larger than 500 kb differentiated the Setaria and rice genomes (Table 8.2).

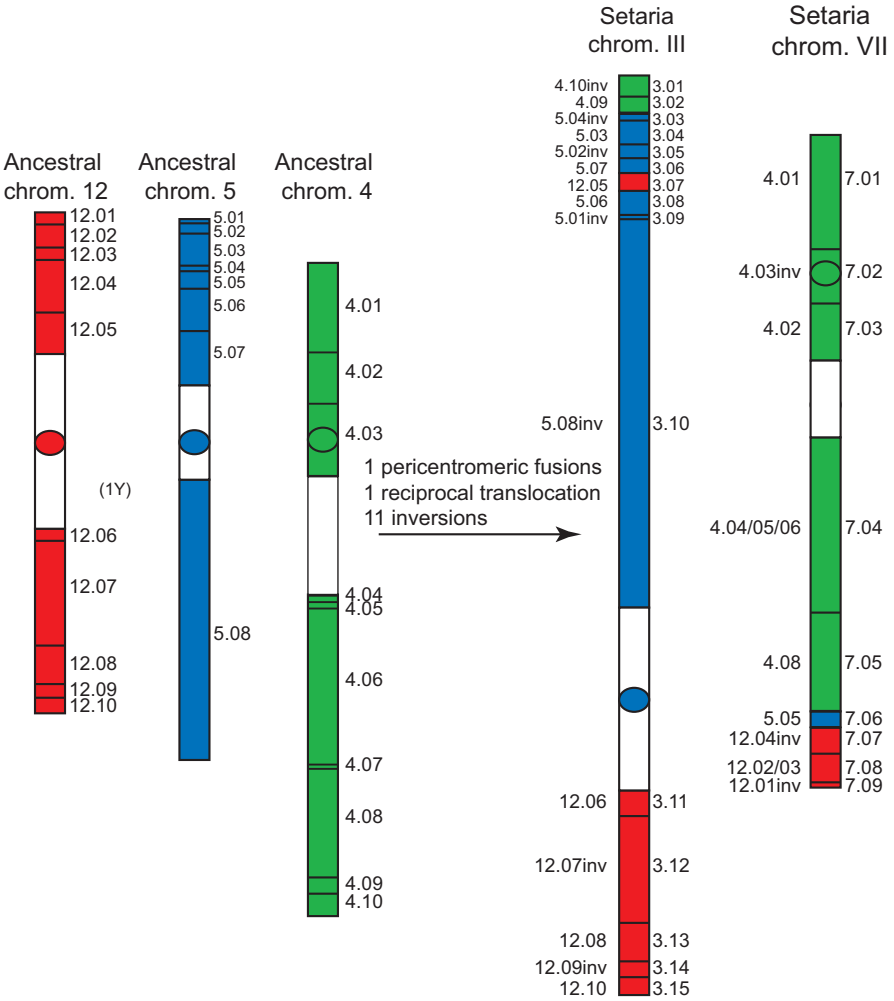


**Fig. 8.2** Structure of three ancestral grass chromosomes and their two syntenic foxtail millet chromosomes. *Circles* indicate centromere positions. *White* areas indicate regions for which little comparative information is available

Extending the comparisons to sorghum ($n=10$) and *Brachypodium* ($n=5$) allowed us to establish the structure of the grass ancestral chromosomes and, consequently, determine how chromosomes evolved to form extant Setaria, sorghum, rice, and *Brachypodium* chromosomes. While it was not always possible to unambiguously determine on which phylogenetic branch a rearrangement took place, we assumed that (1) rearrangements that were differentially present between Setaria/sorghum and rice/*Brachypodium* most likely originated in the lineage leading to sorghum and Setaria, as the two Panicoideae species diverged 26 MYA compared to 52 MYA for the two Pooideae species (Bennetzen et al. 2012); (2) of the four species analyzed,

**Table 8.2** Syntenic blocks in the ancestral grass genome and corresponding blocks and their breakpoints (in parenthesis; in Mb) in the rice, Setaria, and sorghum chromosomes

| Grass ancestor | Rice (Os) | Setaria (Set) | Sorghum (Sb) |
|---|---|---|---|
| 1.01 | 1.01 (0–6.51) | 5.02 (5.62–13.72) | 3.01inv (10.01–1.47) |
| 1.02 | 1.01 (6.52–8.15) | 5.01inv (1.16–0.10) | 3.01inv (1.46–0.03) |
| 1.03 | 1.01 (8.16–12.10) | 5.03inv (5.54–2.18) | 3.02 (10.09–15.84) |
| 1.04 | 1.01 (12.25–21.97) | 5.04 (13.76–24.96) | 3.02 (15.85–49.15) |
| 1.05 | 1.01 (22.00–22.75) | 5.05inv (25.55–24.99) | 3.03inv (50.80–49.27) |
| 1.06 | 1.01 (22.75–23.72) | 5.06 (25.56–26.80) | 3.04 (50.83–52.05) |
| 1.07 | 1.01 (23.97–24.56) | 5.08inv (32.25–31.66)[a] | 3.04 (52.09–53.19) |
| 1.08 | 1.01 (24.58–28.78) | 5.07 (26.85–31.61) | 3.04 (53.19–58.55) |
| 1.09 | 1.01 (28.79–44.94) | 5.09 (32.28–47.22) | 3.04 (58.57–74.44) |
| | 0 rearrangements | 2 Set/Sb inversion; 2 Set inversions | 2 Set/Sb inversion |
| 2.01 | 2.01 (0–0.56) | 1.01 (0–0.49) | 4.01 (0–0.69) |
| 2.02 | 2.01 (0.56–9.67) | 1.02inv (10.39–0.52) | 4.01 (0.71–13.90) |
| 2.03 | 2.01 (9.91–26.54) | 1.03 (10.43–32.52) | 4.01 (13.89–57.12) |
| 2.04 | 2.01 (26.55–32.82) | 1.03 (32.53–38.63) | 4.02inv (63.72–57.15) |
| 2.05 | 2.01 (32.83–36.78) | 1.03 (38.77–42.06) | 4.03 (63.73–68.00) |
| | 0 rearrangements | 1 Set inversion | 1 Sb inversion |
| 3.01 | 3.01 (37.29–36.02) | 9.01 (0.10–1.16) | 1.01 (0.01–1.84) |
| 3.02 | 3.01 (36.01–35.69) | 9.02inv (1.40–1.18) | 1.01 (1.84–2.14) |
| 3.03 | 3.01 (35.70–20.05) | 9.03 (1.46–13.31) | 1.01 (2.14–17.11) |
| 3.04 | 3.01 (19.84–9.22) | 9.07 (41.57–50.13) | 1.03 (54.45–63.18) |
| 3.05 | 3.02 (9.21–8.48) | 9.08inv (50.74–50.15) | 1.04 (63.20–63.82) |
| 3.06 | 3.03 (8.48–7.16) | 9.09 (50.76–51.87) | 1.05 (63.82–65.04) |
| 3.07 | 3.04inv (6.67–7.14) | 9.09 (51.87–52.37) | 1.05 (65.05–65.54) |
| 3.08 | 3.05 (6.64–0.02) | 9.09 (52.37–58.83) | 1.05 (65.55–73.83) |
| 10.01 | 10.01inv (20.60–15.82) | 9.04 (13.32–19.51) | 1.02 (17.14–25.23) |
| 10.01 | 10.01inv (15.26–13.47) | 9.05inv (22.44–20.09)[b] | 1.02 (25.25–31.98) |
| 10.01 | 10.01inv (3.80–0.09) | 9.06 (31.62–35.12) | 1.02 (41.82–47.88) |
| 10.02 | 10.02inv (23.70–20.64) | 9.06 (35.14–41.08) | 1.02 (47.88–53.88) |
| | 1 Os inversion (Jia et al. 2013); 2 Os inversions (Nagamura et al. 1995) | 1 Set/Sb pericentromeric fusion; 2 Set inversions | 1 Set/Sb pericentromeric fusion |
| 4.01 | 4.01 (0.06–4.77) | 7.01 (0.39–5.45) | 6.01 (0.03–3.68) |
| 4.02 | 4.01 (5.25–7.49) | 7.03 (9.39–12.41) | 6.02inv (6.12–3.80) |
| 4.03 | 4.01 (8.30–11.77) | 7.02inv (9.19–6.89) | 6.03 (6.23–15.97) |
| 4.04 | 4.01 (18.39–18.63) | 7.04 (16.62–16.75) | 6.04 (26.36–31.22) |
| 4.05 | 4.01 (18.63–18.93) | 7.04 (16.78–17.42) | 6.05inv (37.25–32.50) |
| 4.06 | 4.01 (19.22–27.72) | 7.04 (17.64–26.33) | 6.06 (37.68–53.27) |
| 4.07 | 4.01 (27.75–27.90) | 2.04inv (22.84–22.69)[c] | 6.06 (53.29–53.49) |
| 4.08 | 4.01 (27.92–33.72) | 7.05 (26.36–31.75) | 6.06 (53.54–59.39) |

(continued)

**Table 8.2** (continued)

| Grass ancestor | Rice (Os) | Setaria (Set) | Sorghum (Sb) |
|---|---|---|---|
| 4.09 | 4.01 (33.89–34.71) | 3.02 (1.25–2.01) | 6.06 (59.50–60.54) |
| 4.10 | 4.01 (34.72–35.96) | 3.01inv (1.23–0) | 6.06 (60.54–62.14) |
| 5.01 | 5.01 (0.04–0.20) | 3.09inv (7.90–7.71) | 9.01 (0.02–0.19) |
| 5.02 | 5.01 (0.21–0.90) | 3.05inv (4.59–3.83) | 9.01 (0.20–1.65) |
| 5.03 | 5.01 (0.92–2.62) | 3.04 (2.49–3.82) | 9.01 (1.80–4.04) |
| 5.04 | 5.01 (2.71–3.03) | 3.03inv (2.42–2.19) | 9.01 (4.06–4.49) |
| 5.05 | 5.01 (3.05–3.85) | 7.06 (31.76–32.46) | 9.01 (4.60–6.17) |
| 5.06 | 5.01 (4.10–6.20) | 3.08 (6.45–7.63) | 9.01 (6.26–9.12) |
| 5.07 | 5.01 (6.29–9.22) | 3.06 (4.59–5.41) | 9.01 (9.60–15.52) |
| 5.08 | 5.01 (14.39–29.83) | 3.10inv (29.34–7.92) | 9.01 (16.33–59.61) |
| 12.01 | 12.01 (0–0.61) | 7.09inv (35.91–35.58) | 8.01 (0.02–0.89) |
| 12.02 | 12.01 (0.97–1.81) | 7.08 (34.13–35.09) | 8.02inv (2.79–1.04) |
| 12.03 | 12.01 (2.29–2.71) | 7.08 (35.13–35.52) | 8.03 (3.18–3.88) |
| 12.04 | 12.01 (2.83–5.46) | 7.07inv (33.95–32.52) | 8.03 (3.94–9.75) |
| 12.05 | 12.01 (5.70–8.01) | 3.07inv (6.43–5.44) | 8.03 (9.83–27.87) |
| 12.06 | 12.01 (17.46–18.14) | 3.11 (39.35–40.74) | 8.03 (34.90–39.01) |
| 12.07 | 12.01 (18.25–23.88) | 3.12inv (46.59–40.88) | 8.03 (39.07–49.72) |
| 12.08 | 12.01 (23.95–26.03) | 3.13 (46.65–48.83) | 8.03 (49.75–52.73) |
| 12.09 | 12.01 (26.04–26.78) | 3.14inv (49.68–48.84) | 8.03 (52.91–53.94) |
| 12.10 | 12.01 (26.82–27.62) | 3.15 (49.74–50.60) | 8.03 (54.04–55.42) |
| | 0 rearrangements | 1 Set pericentromeric fusion; 1 Set reciprocal translocation; 11 Set inversions (8 in chrom. III, 3 in chrom. VII) | 2 Sb inversions (Choi et al. 2004); 0 rearrangements (Wang et al. 1998); 1 Sb inversion (Daverdin et al. 2015) |
| 6.01 | 6.01 (0.16–2.51) | 4.01 (0.11–2.46) | 10.01 (0.01–2.97) |
| 6.02 | 6.01 (2.51–5.04) | 4.02inv (5.84–2.60) | 10.01 (2.98–5.99) |
| 6.03 | 6.01 (5.20–5.96) | 4.03inv (7.20–5.95) | 10.01 (6.21–7.29) |
| 6.04 | 6.01 (6.02–22.25) | 4.04 (7.28–30.14) | 10.01 (7.31–49.38) |
| 6.05 | 6.01 (22.30–26.76) | 4.05inv (36.43–32.42) | 10.01 (49.46–54.08) |
| 6.06 | 6.02inv (28.38–26.80) | 4.05inv (32.34–30.21) | 10.01 (54.15–56.06) |
| 6.07 | 6.03 (28.47–30.55) | 4.06inv (38.61–36.44) | 10.01 (56.10–59.00) |
| 6.08 | 6.03 (30.56–32.06) | 4.07 (38.68–40.22) | 10.01 (58.99–60.97) |
| | 1 Os inversion | 4 Set inversions | 0 rearrangements |
| 7.01 | 7.01 (0–17.04) | 2.01 (0–14.69) | 2.01 (0–18.17) |
| 7.02 | 7.01 (17.40–30.25) | 2.07 (39.99–49.11) | 2.04 (67.44–77.88) |
| 9.01 | 9.01 (0.41–2.71) | 2.03 (16.19–20.12) | 2.03 (22.50–27.81) |
| 9.02 | 9.02inv (5.64–4.09) | 2.03 (20.36–22.62) | 2.03 (27.85–37.82) |
| 9.03 | 9.03 (6.47–7.00) | 2.05inv (24.20–23.43) | 2.03 (45.30–47.38) |
| 9.04 | 9.03 (7.06–23.18) | 2.06 (24.32–39.98) | 2.03 (47.44–67.44) |
| 9.05 | 9.03 (23.26–23.79) | 2.02 (14.90–16.15) | 2.02 (18.59–22.49) |

(continued)

**Table 8.2** (continued)

| Grass ancestor | Rice (Os) | Setaria (Set) | Sorghum (Sb) |
|---|---|---|---|
| | 0 rearrangements (Devos et al. 1998); 1 Os inversion (Wang et al. 1998) | 1 Set/Sb pericentromeric fusion; 2 Set/Sb inversions; 1 Set inversion | 1 Set/Sb pericentromeric fusion; 2 Set/Sb inversion |
| 8.01 | 8.01 (0.07–0.68) | 6.01 (0.15–0.89) | 7.01 (0.07–1.25) |
| 8.02 | 8.01 (0.70–8.50) | 6.02inv (8.06–0.91) | 7.01 (1.26–11.50) |
| 8.03 | 8.01 (8.58–10.45) | 6.03 (8.09–12.31) | 7.01 (11.66–22.59) |
| 8.04 | 8.01 (10.57–11.10) | 6.03 (12.64–14.52) | 7.03 (39.67–47.50) |
| 8.05 | 8.01 (15.81–17.07) | 6.03 (20.65–21.62) | 7.02inv (39.08–36.76) |
| 8.06 | 8.01 (17.67–18.19) | 6.04inv (22.73–21.78) | 7.04 (48.37–49.91) |
| 8.07 | 8.01 (18.28–23.80) | 6.05 (23.65–31.07) | 7.04 (50.01–58.34) |
| 8.08 | 8.01 (23.97–28.49) | 6.05 (31.12–35.97) | 7.05inv (64.31–58.36) |
| | 0 rearrangements | 2 Set inversions | 3 Sb inversions |
| 11.01 | 11.01 (0.04–0.19) | 8.01 (0.07–0.17) | 5.01 (0.05–0.20) |
| 11.02 | 11.01 (0.23–1.60) | 8.02inv (1.14–0.21) | 5.01 (0.28–2.21) |
| 11.03 | 11.01 (1.65–5.38) | 8.03 (1.22–7.87) | 5.01 (2.25–10.62) |
| 11.04 | 11.01 (5.44–9.01) | 8.04inv (13.45–8.01) | 5.02inv (19.22–10.63) |
| 11.05 | 11.01 (9.03–31.17) | 8.05 (13.57–40.59) | 5.03 (19.54–62.33) |
| | 0 rearrangements | 1 Set/Sb inversion; 1 Set inversion | 1 Set/Sb inversion |

[a]Block 5.08 is likely misplaced and misoriented in the Yugu1 assembly and should be located between 5.06 and 5.07 (see Table 8.1). This assembly error was not counted as a rearrangement
[b]Yugu-1 specific inversion; this inversion was not considered as an evolutionary rearrangement
[c]The 2.04 segment was placed in this location on the basis of a single SNP marker. It is unclear whether this is a rearrangement or an assembly error, and hence was not considered in our analysis

rice has the most stable genome (Bennetzen and Ma 2003; Massa et al. 2011); and (3) the ancestral genome comprised 12 chromosomes with structures similar to those of extant rice (Devos 2010). The syntenic blocks in the grass ancestral chromosomes, and the corresponding blocks and their breakpoints in Setaria, sorghum, and rice are presented in Table 8.2. A syntenic block in the ancestral grass genome is defined as a chromosomal segment that is colinear in rice, Setaria, and sorghum. In the three extant species, a syntenic block is a contiguous segment in which gene orders have remained conserved following divergence from the ancestral grass genome. For inversions, only rearranged segments larger than 500 kb are listed. Also given in Table 8.2 is the number of rearrangements that was needed to form current-day Setaria, sorghum, and rice chromosomes. Not included in this tally are the two inversions that were likely caused by the misplacement and misorientation of synteny block 5_08 in the Yugu1 sequence (Table 8.2).

Chromosomes III and VII have the most complex structure, and the evolutionary history of these two chromosomes was modeled by Daverdin et al. (2015). Daverdin and colleagues showed that both chromosomes evolved from the same three ancestral

chromosomes and that a minimum of 13 rearrangements was needed to explain their structure, including one pericentromeric fusion, one reciprocal translocation, and 11 inversions. Overall, the nine Setaria chromosomes were formed from 12 ancestral chromosomes following three pericentromeric fusions, one reciprocal translocation, and 29 inversions. Of these, two pericentromeric fusions and five inversions occurred before the divergence of Setaria from a common ancestor with sorghum (Table 8.2). The third pericentromeric fusion was dated to 26–13.1 MYA, that is, after the divergence of the sorghum and Setaria/switchgrass lineages but before the divergence of Setaria and switchgrass (Daverdin et al. 2015). Of the 24 inversions that occurred in the Setaria lineage after its divergence from sorghum, at least three took place before the divergence of Setaria and switchgrass and at least eight after the divergence of these two species. The timing of the remaining inversions is unknown because comparisons with switchgrass were based on genetic map data and hence sufficient comparative data points were available only for regions that spanned at least 1.2 Mb in foxtail millet (Daverdin et al. 2015). The reciprocal translocation was specific to Setaria.

Reconstruction of the evolutionary history of the sorghum and Setaria genomes allowed us to draw conclusions on the relative stability of these two Panicoid genomes during the past 26 MYA after their divergence from a common ancestor. During that time period, the Setaria lineage underwent one pericentromeric fusion, one reciprocal translocation and 24 inversions of segments larger than 500 kb. During the same time frame, the sorghum lineage underwent seven inversions. The same trend was also seen for small inversions that encompassed only a single or a few genes (Bennetzen et al. 2012). A correlation between the rates with which small scale rearrangements and chromosome level rearrangements take place has previously been noted in comparisons between *Ae. tauschii* and other grass species (Massa et al. 2011). It is also known that different lineages can accumulate and fix rearrangements at different rates although what causes differential stability of genomes has not been elucidated (Devos and Gale 2000). The Setaria genome is smaller, and hence has less repeats, than the sorghum genome indicating that repeat content per se is not the driver behind the rearrangements. Or, if rearrangements are associated with the repeat content, another factor must play a role in their fixation. Setaria is a largely selfing species, which decreases the effective population size and increases the importance of drift as a mechanism for fixing rearrangements. However, this is also true for sorghum. There is some evidence that inversions might be involved in adaptation. For example, frequencies of particular inversions in *Drosophila* and *Anopheles* species have been shown to be correlated with environmental gradients (Anderson et al. 2005; Coluzzi et al. 2002). Annual and perennial forms of *Mimulus gutatus* adapted to dry and moist cool areas, respectively, also vary by the presence of an inversion that might have been fixed in one population group as a result of local adaptation (Kirkpatrick 2010). It is conceivable that at least some of the evolutionary inversions were fixed because of differential adaptation of the sorghum and Setaria lineages after their radiation from a common ancestor.

Further insights into the role that local adaptation plays in the fixation of rearrangements might be obtained by conducting comparative analyses at the genome level of large numbers of accessions. Whole genome resequencing data for hundreds of

accessions is becoming available for both *S. italica* and its wild progenitor *S. viridis*. We know that Yugu1 carries a ~2.9 Mb inversion which is absent from Zhang gu as well as from B100 and A10, the *S. italica* and *S. viridis* parents, respectively, used to construct the genetic map to which the Yugu1 sequence was anchored. If this inversion or other rearrangements are found in multiple accessions, it will be interesting to establish whether their presence is associated with specific environmental conditions.

**Acknowledgements** This work was supported by the National Institute of Food and Agriculture Plant Feedstock Genomics for Bioenergy Program (grant no. 2008-35504-04851) and the Office of Biological and Environmental Research in the Department of Energy (DOE) Office of Science to the BioEnergy Science Center (BESC).

## References

Anderson AR, Hoffmann AA, McKechnie SW, Umina PA, Weeks AR. The latitudinal cline in the In(3R)Payne inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations. Mol Ecol. 2005;14(3):851–8.

Bennetzen JL, Freeling M. The unified grass genome: synergy in synteny. Genome Res. 1997;7: 301–6.

Bennetzen JL, Ma J. The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. Curr Opin Plant Biol. 2003;6:128–33.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30:555–61.

Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, et al. Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci. 2004;101:15289–94.

Coluzzi M, Sabatini A, della Torre A, Di Deco MA, Petrarca V. A polytene chromosome analysis of the *Anopheles gambiae* species complex. Science. 2002;298(5597):1415–8.

Daverdin G, Bahri BA, Wu X, Serba DD, Tobias C, Saha MC, et al. Comparative relationships and chromosome evolution in switchgrass (*Panicum virgatum*) and its genomic model, foxtail millet (*Setaria italica*). Bioenerg Res. 2015;8:137–51.

Devos KM. Grass genome organization and evolution. Curr Opin Plant Biol. 2010;13:139–45.

Devos KM, Gale MD. Genome relationships: the grass model in current research. Plant Cell. 2000;12:637–46.

Devos KM, Wang ZM, Beales J, Sasaki T, Gale MD. Comparative genetic maps of foxtail millet (*Setaria italica*) and rice (*Oryza sativa*). Theor Appl Genet. 1998;96:63–8.

Gale MD, Devos KM. Plant comparative genetics after 10 years. Science. 1998;282:656–9.

Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013;45(8):957–61.

Kirkpatrick M. How and why chromosome inversions evolve. PLoS Biol. 2010;8(9):e1000501.

Massa AN, Wanjugi H, Deal KR, O'Brien KO, You FM, Maiti R, et al. Gene space dynamics during the evolution of *Aegilops tauschii*, *Brachypodium distachyon*, *Oryza sativa*, and *Sorghum bicolor* genomes. Mol Biol Evol. 2011;28:2537–47.

Nagamura Y, Inoue T, Antonio BA, Shimano T, Kajiya H, Shomura A, et al. Conservation of duplicated segments between rice chromosomes 11 and 12. Breed Sci. 1995;45:373–6.

Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD. Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. Theor Appl Genet. 1998;96:31–6.

Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, et al. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol. 2012;30: 549–54.

# Chapter 9
# LTR Retrotransposon Dynamics and Specificity in *Setaria italica*

**Jeffrey L. Bennetzen, Minkyu Park, Hao Wang, and Hongye Zhou**

**Abstract** The distributions of different LTR retrotransposon families and structures were analyzed across the ~400 Mb assembly for the ~500 Mb genome of *Setaria italica*. The results indicated different genomic distributions for all five of the highly abundant LTR retrotransposon families that were investigated. Unequal recombination and illegitimate recombination appeared to be more active in LTR retrotransposon removal in the gene-rich regions towards the ends of all chromosomes. In striking contrast to this result, LTR retrotransposon ages did not differ dramatically across the assembled genome, suggesting that LTR retrotransposon removal rates are not dramatically influenced by genomic location. These two, largely incompatible, observations indicate that the dynamics of LTR retrotransposon activation, insertion, and removal all need a great deal of additional investigation, including highly detailed intraspecies analyses and interspecies comparisons.

**Keywords** *Setaria italica* • Foxtail millet • LTR retrotransposons • Unequal recombination • Illegitimate recombination • Insertion times

## 9.1 Introduction

Transposable elements (TEs) are highly abundant in the nuclear genomes of all higher plants, usually constituting the majority of DNA in any species with a genome size exceeding 800 Mb (Bennetzen and Wang 2014). There is continued debate regarding the possible roles of these TEs. Although pure Darwinian theory indicates that selfish sequences like TEs would be obligated to come into existence via the process of natural selection for superior transmission (Doolittle and Sapienza 1980; Orgel and Crick 1980), there are many cases where these TEs have been co-opted by their hosts for a novel function, particularly in gene regulation and sometimes in the creation of new genes (Naito et al. 2009, reviewed in Feschotte 2008

J.L. Bennetzen (✉) • M. Park • H. Wang • H. Zhou
Department of Genetics, University of Georgia, Athens, GA 30677, USA
e-mail: maize@uga.edu



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_9

and in Bennetzen and Wang 2014). Overall, however, the biology and distribution of TEs is consistent with the properties of a selfish DNA (Daniels et al. 1990; Baucom et al. 2009).

In many genomes, including most or all flowering plant genomes (Grandbastien and Casacuberta 2012; Bennetzen and Wang 2014), TEs are the major source of genome rearrangement, either by chromosome breakage, by gene acquisition/mobilization or as sites for ectopic/unequal homologous recombination. In plant species with a strong record of recent TE activity, like maize, more than half of the nuclear genome can be structurally rearranged in as little as 1–2 million years, by a combination of macro and micro events (Wang and Dooner 2006; Wang and Bennetzen 2012).

Despite their ubiquity and importance in plants, there is surprisingly little information regarding the genomic properties or specificities of TEs. TEs within many angiosperm genomes have been detected, although pure whole genome shotgun (WGS) sequences are often masked for repeats, thence leading to a genome assembly that is deficient in most or all of the highly abundant TEs (e.g., Al-Dous et al. 2011). Once the TEs in a sequenced genome are found, the standard next step is for TEs to be broadly categorized by structural or homology criteria (Wicker et al. 2007). However, these searches often involve only discovery of highly repeated elements (which often make up a tiny minority of the TE families in a genome) or elements with homology to TEs that have already been described in other species. Even with the current wealth of deeply sequenced plant genomes (>95), these two criteria still can miss >50 % of the LTR retrotransposon families that are present in a newly sequenced genome (H. Wang and J. Bennetzen, unpub. obs.). Applying sensitive structural criteria in genome analysis and annotation has led to particularly comprehensive TE discovery results (Schnable et al. 2009; Bennetzen et al. 2012; Hellsten et al. 2013).

Once discovered, the TEs in any "fully sequenced" genome can be mapped across chromosomes. This analysis has been undertaken on many occasions and has yielded the routine observation that different classes of TEs show very different abundances and very different genomic distributions. The LTR retrotransposons are routinely the most abundant TEs in plants, and the most routine LTR retrotransposon pattern has been that elements of the *Gypsy* superfamily mostly accumulate in heterochromatin (particularly pericentromeric heterochromatin) and that elements of the *Copia* superfamily show less of a bias towards the pericentromeric regions. More detailed analysis in maize, at the LTR retrotransposon family level, indicates that the likelihood of an LTR retrotransposon family's insertion into euchromatin is inversely proportional to its copy number, regardless of the superfamily designation (Baucom et al. 2009). The DNA TEs, like the classic "controlling element" *Ac/Ds* studied by McClintock (McClintock 1956), tend to exist in lower copy numbers and show preferential association with genes. Of all TE families, the CACTA TEs (for instance, *Spm/dspm* of maize) seem to show the most "random" distribution, but all evidence to date for eukaryotes and prokaryotes indicates that TEs are far from random in their insertion or accumulation specificities (Schnable et al. 2009; Bennetzen et al. 2012).

The very complete sequence of the *Setaria italica* genome (Bennetzen et al. 2012) provides a particularly useful resource for TE characterization. In the first analyses, using the full spectrum of search criteria (repetitiveness, homologies to

known elements, and structural properties), TEs were found to constitute at least 40 % of the genome. Most of the TE genome space (>60 %) comprises LTR retrotransposons, including 98, 107, and 361 families of *Gypsy*, *Copia,* and "superfamily-unknown" LTR retrotransposons. As seen with many previously studied plant genomes, the *Gypsy* LTR retrotransposons were found to be enriched in the pericentromeric heterochromatin (Bennetzen et al. 2012). However, this general *Gypsy* observation might be driven by the properties of a few very abundant families, so a more detailed family-by-family analysis is warranted. This chapter provides that analysis and also includes investigations of the processes, rate, and genomic specificities of the loss of LTR retrotransposon sequences from the *S. italica* genome.

## 9.2 Results

### *9.2.1 The Distributions of Different LTR Retrotransposon Families*

LTR retrotransposon families within a species are designated as distinct by an "80 % homology rule" for their LTR nucleotide sequences (Wicker et al. 2007). This threshold was not purely arbitrary in its choice because the homology for LTR retrotransposons that show close internal relatedness is usually quite high (>90 %) while those that lack close internal relatedness show little LTR homology (<50 %). LTRs were chosen as the defining sequence source because they are more commonly intact (exclusively in solo LTRs, but also in complete/intact elements and fragments) than are the internal coding sequences. Figure 9.1 shows the distribution of the five most abundant families of LTR retrotransposons from the *Gypsy* and *Copia* superfamilies for two scaffolds in *S. italica*. As is obvious, the distributions of these families are quite different. Because all LTR retrotransposons appear to be removed from genomes by the same processes (Devos et al. 2002, reviewed in Bennetzen and Wang 2014), the differences between families are most likely to be caused by differences in insertion preferences.

#### 9.2.1.1 LTR Retrotransposon Removal

LTR retrotransposons have been defined by our lab as intact (having both LTRs, and the appropriate target site duplications (TSDs)), fragmented (missing at least all or part of one LTR and usually missing some internal sequences), or solo LTRs (an LTR with no internal sequences, but with a TSD indicating that the solo LTR originated from unequal recombination between the two LTRs). Transposition of an LTR retrotransposon requires that it be intact, with two LTRs that are of the appropriate structure and orientation. An intact LTR retrotransposon does not need to have all of its internal coding sequences functional because some of the transposition functions

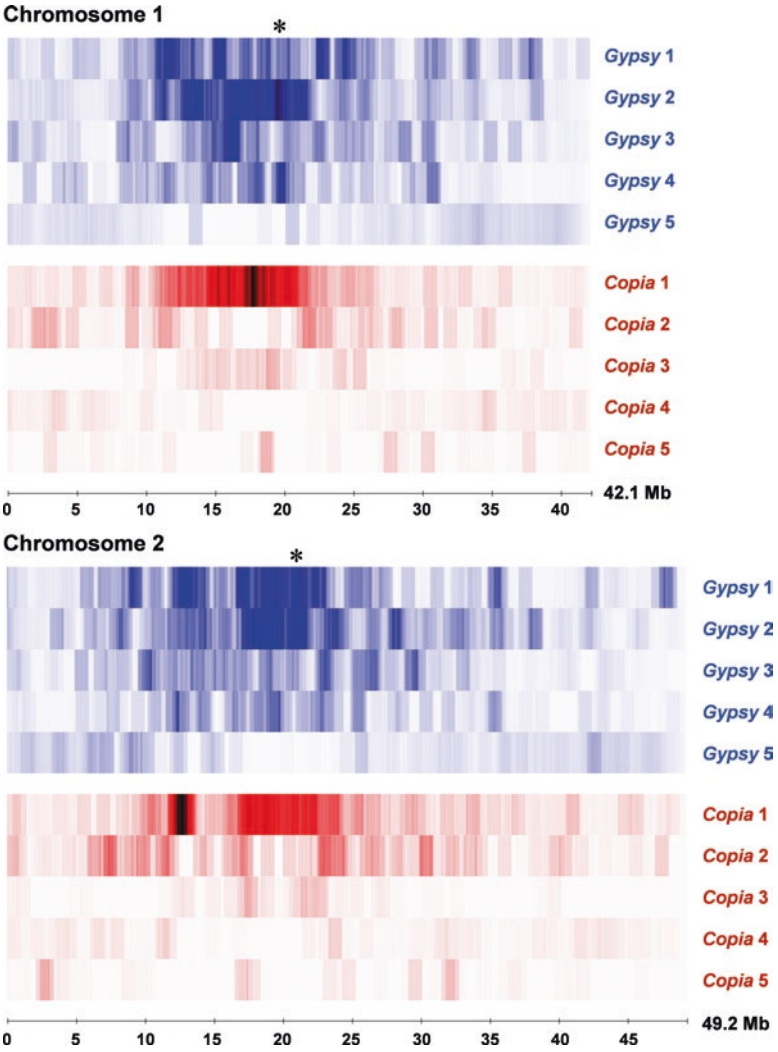


**Fig. 9.1** Distributions of the ten most abundant LTR retrotransposon families across two *S. italica* chromosomes. The heat map was derived from a sliding window analysis of 1 Mb each, with 10 kb steps. Higher pigment density indicates higher LTR retrotransposon density, as determined by intact LTR retrotransposon amount (kb) in that window. The * indicates the approximate position of the centromere (Bennetzen et al. 2012). This same analysis was performed for all nine *S. italica* chromosomes, with similar results, but only two are shown here due to space considerations. It should be noted that these are actually scaffold depictions rather than full chromosomes, with numerous sequence gaps, especially in the centromeric regions

can be provided in trans (Jin and Bennetzen 1989), but it does need to have a primer binding site (PBS) and a polypurine tract (PPT), two short internal sequences needed for synthesis of the two strands of the integration intermediate. Because only an intact LTR retrotransposon can make an appropriate copy for transposition, then all solo LTRs and fragmented elements are an indication of LTR retrotransposon decay and removal (Devos et al. 2002). Figure 9.2 shows the distributions of solo LTRs, fragmented LTR retrotransposons, and intact LTR retrotransposons across two *S. italica* scaffolds.

As noted previously in rice (Ma and Bennetzen 2006), solo LTRs are in relatively low abundance in the low recombination regions around the centromere (*). This is an expected outcome of the low level of homologous recombination in these regions, and thus underscores the very rapid rate at which solo LTRs can be generated in euchromatic regions by unequal recombination. The fragmented LTR retrotransposons, primarily derived from various deletion processes (Kirik et al. 2000; Devos et al. 2002; Wicker et al. 2010), are also found primarily where LTR retrotransposons of all levels of intactness are most abundant (Fig. 9.2). Surprisingly, the fragmented LTR retrotransposons also show a higher ratio to intact LTR retrotransposons in the distal regions of chromosome arms. This indicates that the deletion mecha-

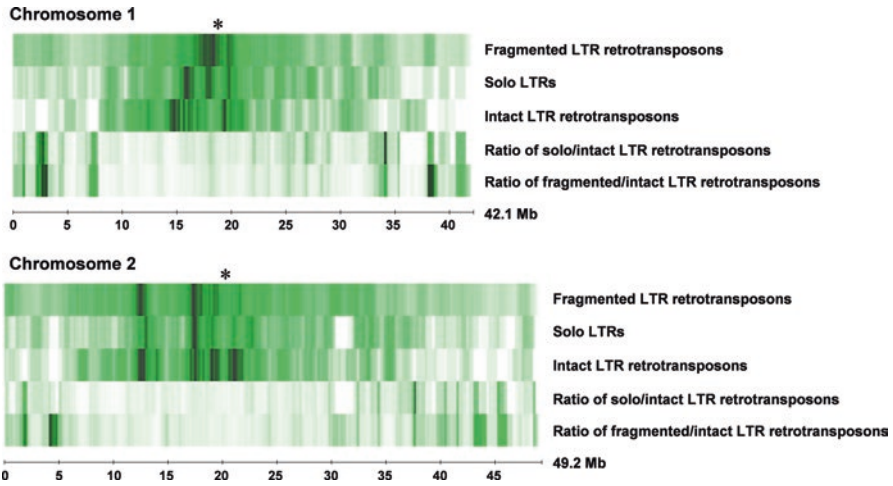


**Fig. 9.2** Distributions and ratios of solo LTRs, fragmented LTR retrotransposons, and intact LTR retrotransposons across two *S. italica* chromosomes. The heat map was derived from a sliding window analysis of 1 Mb, with 100 kb steps. Higher pigment density indicates higher element density, as measured in kb in that window. In order to smooth the ratio curves, two kb of intact LTR retrotransposon was added to each window so that the denominator in the ratio was never zero. Hence, for this and several other reasons, the actual values of these ratios are not important, but the difference in the ratios should be noted across the chromosomal length. "Fragmented" LTR retrotransposons are defined as those that lack at least part of one LTR (and, often, additional internal sequences), while "intact" LTR retrotransposons are defined as those with two appropriate LTRs (in correct orientation and full size). The * indicates the approximate position of the centromere. This same analysis was performed for all nine *S. italica* chromosomes, with similar results, but only two are shown here due to space considerations

nisms involved in their fragmentation are not evenly active across the genome. This is an expected result for unequal recombination (which is known to preferentially occur in euchromatin) but not necessarily for fragmentation caused by illegitimate recombination (where possible differences in rates across the genome have not yet been thoroughly investigated).

### *9.2.2 Insertion Times of Detected LTR Retrotransposons*

At the time of insertion, the mechanism of LTR retrotransposon replication indicates that the two LTRs in a single element will almost always be identical in sequence. Hence, divergence of the two LTRs in an element can be used to date the time that has elapsed since insertion (SanMiguel et al. 1998). This molecular clock has not been calibrated nearly so well as the molecular clock for "neutral" third codon positions in genes and is also likely to be somewhat variable between chromosome locations and species. Still, it provides an estimate of insertion date that is likely to be fairly consistent within a single genome for LTR retrotransposons with similar insertion biases. Unexpectedly, the results depicted in Fig. 9.3 indicate relatively similar average ages of LTR retrotransposons across the Setaria genome, suggesting that removal of TEs is not much more rapid from euchromatin than it is from heterochromatin.

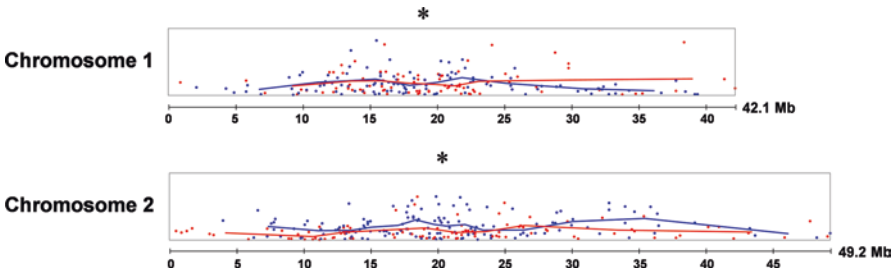


**Fig. 9.3** Approximate insertion dates of all intact LTR retrotransposons across two *S. italica* chromosomes. *Individual dots* indicate individual elements and their insertion sites. The *curves* indicate average ages for the LTR retrotransposons of either the *Gypsy* (*red*) or *Copia* (*blue*) superfamilies in each 10-element window. The * indicates the approximate position of the centromere. This same analysis was performed for all nine *S. italica* chromosomes, with similar results, but only two are shown here due to space considerations. The *X* axis indicates position along the chromosome, and the *Y* axis indicates the degree of sequence divergence between two LTRs in the same element, which is an indication of the time that has expired since the TE inserted. Larger values mean more ancient insertion dates

## 9.3 Discussion

Although LTR retrotransposons make up the majority of most plant genomes, their distributions, origins, specificities, and transposition histories are only crudely understood in any plant system. More research is needed to look across a broad (and informatively selected) phylogenetic spectrum of organisms to see how TEs are transmitted and behave. In plants, where the median genome size is >4 Gb, the cost of de novo investigation of this process across hundreds of taxa remains prohibitive if one requires a fully sequenced and annotated genome. However, sample sequencing followed by repeat annotation of genomes can allow a cost-effective approach for this analysis (Devos et al. 2005; Macas et al. 2007). For instance, using only a few thousand Sanger sequences, we discovered that the doubling of the *Zea luxurians* genome size in the last 1–3 million years was caused by the amplification of numerous (but not all) LTR retrotransposon families in that genomic lineage (Estep et al. 2013).

Even when a genome is fully sequenced, the TEs are often given only a cursory investigation. This was certainly true in our earlier sequencing and annotation of the *S. italica* genome (Bennetzen et al. 2012). Even the much more comprehensive analysis of the TEs in the maize genome (Schnable et al. 2009; Baucom et al. 2009; Yang and Bennetzen 2009) left numerous questions regarding ancestry, activity, specificity, and fate unanswered. Questions investigated in detail in maize (but mostly not in any other genome sequence description in plants) were the number of LTR retrotransposon families (406), the frequency of gene fragment acquisition (high for *Helitrons* but also seen for MITEs, LTR retrotransposons, and other TE types), and TE distributions across the chromosomes. In *S. italica*, relatively little LTR retrotransposon data analysis was presented although >500 families were identified and mapped across the genome (Bennetzen et al. 2012 and unpub. data).

In our current analysis, we show that different LTR retrotransposon families have very different genomic distributions, as noted previously in maize (Baucom et al. 2009). It has been proposed (Bennetzen 2000, 2005) that this apparent differential insertion specificity is caused by recognition of different chromatin states, as is known to be the case for LTR retrotransposons in yeast (Kirchner et al. 1995; Zou and Voytas 1997). As future studies investigate chromatin structure in Setaria and other plants in more detail, it will be interesting to see which associations hold up between domains in integrase (the TE enzyme involved in opening up the host DNA for LTR retrotransposon insertion) and specific chromatin compositions and/or configurations. For instance, LTR retrotransposons that encode a chromodomain are more likely to be found in heterochromatic regions (Gao et al. 2008), suggesting one initial level of very general specificity. However, as there are likely to be thousands of actual "types" of heterochromatin or euchromatin, sporting different protein modifications and compositions (Bennetzen 2000), it is likely that integrases will have sufficient diversity to find unique target sites for most or all LTR retrotransposon families.

Although the general story that LTR retrotransposons in plants preferentially accumulate in pericentromeric regions has been adopted as a general concept, it is not a particularly accurate representation of reality. As seen previously in maize

(Baucom et al. 2009), this pericentromeric bias is quite common for the most abundant LTR retrotransposons, but has dramatic exceptions (see, for instance, the second most abundant *Copia* family in Setaria, *Copia 2*, which strongly avoids pericentromeric heterochromatin). Exceptions are also found to the direct relationship between copy number of the LTR retrotransposon and its preferential accumulation in heterochromatin although that role does hold true on average. It will be interesting, in future studies, to see what insertion niche is found by high-copy-number LTR retrotransposon families that do not find a safe haven in heterochromatin, as seen for *Gypsy 2*, *Copia 2*, and *Copia 4* in this study. We predict that these elements will not be found inserted into genes, but in class(es) of small heterochromatic blocks that are interspersed with genes, or perhaps in gene regulatory regions. If the latter, then these would be an excellent set of TEs to investigate for their ability to bring new genetic diversity to gene regulation (Kidwell and Lisch 1997; Feschotte 2008).

Our two most surprising results were (1) the lack of any dramatic difference in average LTR retrotransposon age between pericentromeric and euchromatic regions of *S. italica* chromosomes and (2) very uneven ratios of fragmented LTR retrotransposons to intact retrotransposons across the chromosomes. LTR retrotransposons were seen to average a somewhat more ancient time of insertion in pericentromeric heterochromatin, but a stronger effect was expected because both natural selection (to remove TEs that cause mutations) and random ectopic recombination are expected to decrease LTR content especially rapidly in genic regions. The relatively low frequency of solo LTRs in pericentromeric heterochromatin that we observed agrees with this prediction. Perhaps our results are caused by the fact that the LTR retrotransposons at the ends of the chromosomes are so rare that a useful ratio could not be determined or by the fact that not all of the pericentromeric DNA was assembled for *S. italica*, due to its highly repetitive DNA content. If values for these most dissimilar regions had been plotted, then one expects that a more impressive differential would have been observed. The second issue remains even more mysterious. One model suggests that the major mode of DNA removal from plants involves small deletion caused by illegitimate recombination, primarily as an outcome of inaccurate double strand break repair (DSBR) (reviewed in Bennetzen 2007). It is not at all clear why DSBR would be less common or less accurate in heterochromatic regions, but this is certainly implied by our data. More comprehensive analyses are needed investigating specific sequence change types and rates across plant chromosomes.

As with every other TE study conducted in plants, one is inundated with enormous numbers of possible interspecies and intraspecies investigations. Such studies can focus on TE effects on genome structure, on genome function and/or on gene evolution. Our study has provided a small part of this analysis, finding general similarity and some interesting differences with comparable studies in maize, rice, and other angiosperm genomes. We look forward to future studies that will investigate additional properties of these dynamic genome components.

**Acknowledgements** These analyses and the writing of this manuscript were supported by funding from the Giles Professorship to J. L. B. We thank Aye Htun for assistance with manuscript production.

## References

Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, et al. *De novo* genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). Nat Biotechnol. 2011;29:521–7.

Baucom RS, Estill JC, Chaparro C, Upshaw N, Jogi A, Deragon JM, et al. Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. PLoS Genet. 2009;5:e1000732.

Bennetzen JL. The many hues of plant heterochromatin. Genome Biol. 2000;1:Reviews107.

Bennetzen JL. Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev. 2005;15:621–7.

Bennetzen JL. Patterns in grass genome evolution. Curr Opin Plant Biol. 2007;10:176–81.

Bennetzen JL, Wang H. The contributions of transposable elements to the structure, function, and evolution of plant genomes. Annu Rev Plant Biol. 2014;65:505–30.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant *Setaria*. Nat Biotechnol. 2012;30:555–61.

Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A. Evidence for horizontal transmission of the P-transposable elements between Drosophila species. Genetics. 1990;124:339–55.

Devos KM, Brown JKM, Bennetzen JL. Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. Genome Res. 2002;12:1075–9.

Devos KM, Ma J, Pontaroli AC, Pratt LH, Bennetzen JL. Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. Proc Natl Acad Sci U S A. 2005;102:19243–8.

Doolittle WF, Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. Nature. 1980;284:601–3.

Estep MC, DeBarry JD, Bennetzen JL. The dynamics of LTR retrotransposon accumulation across 25 million years of panicoid grass evolution. Heredity. 2013;110:194–204.

Feschotte C. Opinion—transposable elements and the evolution of regulatory networks. Nat Rev Genet. 2008;9:397–405.

Gao X, Hou Y, Ebina H, Levin HL, Voytas DF. Chromodomains direct integration of retrotransposons to heterochromatin. Genome Res. 2008;18:359–69.

Grandbastien M-A, Casacuberta JM. Plant transposable elements impact on genome structure and function. Berlin: Springer; 2012. SpringerLink (Online service). doi:10.1007/978-3-642-31842-9.

Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, Wessler SR, et al. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. Proc Natl Acad Sci U S A. 2013;110:19478–82.

Jin YK, Bennetzen JL. Structure and coding properties of Bs1, a maize retrovirus-like transposon. Proc Natl Acad Sci U S A. 1989;86:6235–9.

Kidwell MG, Lisch D. Transposable elements as sources of variation in animals and plants. Proc Natl Acad Sci U S A. 1997;94:7704–11.

Kirchner J, Connolly CM, Sandmeyer SB. Requirement of RNA polymerase III transcription factors for *in vitro* position-specific integration of a retroviruslike element. Science. 1995;267:1488–91.

Kirik A, Salomon S, Puchta H. Species-specific double-strand break repair and genome evolution in plants. EMBO J. 2000;19:5562–6.

Ma J, Bennetzen JL. Recombination, rearrangement, reshuffling, and divergence in a centromeric region of rice. Proc Natl Acad Sci U S A. 2006;103:383–8.

Macas J, Neumann P, Navratilova A. Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. BMC Genomics. 2007;8:427.

McClintock B. Controlling elements and the gene. Cold Spring Harb Symp Quant Biol. 1956;21:197–216.

Naito K, Zhang F, Tsukiyama T, Saito H, Hancock CN, Richardson AO, et al. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature. 2009;461:1130–4.

Orgel LE, Crick FH. Selfish DNA: the ultimate parasite. Nature. 1980;284:604–7.

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL. The paleontology of intergene retrotransposons of maize. Nat Genet. 1998;20:43–5.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, et al. The B73 maize genome: complexity, diversity, and dynamics. Science. 2009;326:1112–5.

Wang H, Bennetzen JL. Centromere retention and loss during the descent of maize from a tetraploid ancestor. Proc Natl Acad Sci U S A. 2012;109(51):21004–9.

Wang QH, Dooner HK. Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. Proc Natl Acad Sci U S A. 2006;103:17644–9.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 2007;8:973–82.

Wicker T, Buchmann JP, Keller B. Patching gaps in plant genomes results in gene movement and erosion of colinearity. Genome Res. 2010;20:1229–37.

Yang L, Bennetzen JL. Distribution, diversity, evolution, and survival of *Helitrons* in the maize genome. Proc Natl Acad Sci U S A. 2009;106:19922–7.

Zou S, Voytas DF. Silent chromatin determines target preference of the Saccharomyces retrotransposon Ty5. Proc Natl Acad Sci U S A. 1997;94:7412–6.

# Part III
# Morphology

# Chapter 10
# Morphological Development of *Setaria viridis* from Germination to Flowering

**John G. Hodge and Andrew N. Doust**

**Abstract** The model system *Setaria viridis* is morphologically similar to other members of the Panicoideae, including maize and sorghum, although as a wild lineage it still contains a great deal of developmental plasticity. Underlying this variation is a robust ontogenetic pattern of vegetative growth resulting in the production of semi-independent basal branches (tillers) in addition to aerial branches on the main culm and tillers. We characterize the life cycle of *S. viridis* from germination to flowering and the general patterns of vegetative growth that can be expected within this period when grown under standardized conditions. We also indicate what can be expected when these plants are grown under other conditions.

**Keywords** *Setaria viridis* • Green foxtail • Vegetative morphology • Tillering • Branching • Flowering

## 10.1 Introduction

The panicoid grasses comprise approximately 3240 species (Kellogg 2001), distributed worldwide in temperate and tropical ecosystems. Several are dominants in warm-temperate prairie ecosystems, and different species have been domesticated as cereal grains in various parts of the world, including maize in central Mexico (*Zea mays*), sorghum in sub-Saharan Africa (*Sorghum bicolor*), pearl millet in southern Africa (*Pennisetum glaucum*), and foxtail millet in northern China (*Setaria italica*). The rapid growth and biomass accumulation of some wild species has also marked them as potential biofuel sources, including switchgrass (*Panicum virgatum*) and Miscanthus (*Miscanthus* x *giganteus*). Despite the broad agricultural and ecological importance of these grasses, basic research has been primarily focused within the domesticated cereals, with a strong focus on maize. The sexual dimorphism of terminal staminate tassel inflorescences and axillary pistillate ear

J.G. Hodge (✉) • A.N. Doust
Department of Plant Biology, Ecology, and Evolution, Oklahoma State University, 301 Physical Science, Stillwater, OK 74078, USA
e-mail: jghodge@okstate.edu



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_10

inflorescences in maize allows for easy and elegant genetic manipulation, but further functional characterization can be laborious compared to other plant models like *Arabidopsis thaliana*, given the large size of individual plants, polyploid genome, a life cycle that can take over 2 months, and difficulty of transformation (Ishida et al. 2007). In addition, maize and other domesticated panicoid cereals have highly derived phenotypes resulting from human mediated selection (Abbo et al. 2014).

Recently, Setaria has been suggested as a system that incorporates both a domesticated species, foxtail millet (*Setaria italica*) and its wild progenitor, green foxtail (*S. viridis*). In green foxtail, the highly inbred accession line A10.1 is being actively developed as a model for functional and developmental research [(Brutnell et al. 2010), Chaps 10, 11, 13, 14, 17–21]. *S. viridis* A10.1 (hereafter referred to simply as *S. viridis*) has many advantages that make this system more tractable as a model than maize and other agricultural counterparts. Among them are its smaller size, often only reaching 30–40 cm in height at maturity, small genome, and rapid life cycle with heading occurring roughly 21 days after germination, followed by seed maturation of the primary inflorescence within ~40 days post-germination. Alongside the capacity for high efficiency transformation through tissue culture [Chap. 20] more rapid transformation methods of *S. viridis* using floral dip protocols may also be possible [(Martins et al. 2015), (Van Eck and Swartwood 2014), Chap. 21]. *S. viridis* also benefits from a moderately large seed set (each inflorescence often bearing more than 100 seeds), allowing for easy bulking within each generation.

In this chapter, we describe the general development of *S. viridis* from germination to flowering that can be expected when Setaria is grown under standardized conditions. We focus primarily on vegetative development as previous work has described inflorescence development in the genus (Doust and Kellogg 2002; Doust et al. 2005). We also highlight the developmental lability of *S. viridis,* as it, along with other wild grasses, retains the capacity to recognize and respond dynamically to environmental cues during development, and can display a wide array of growth forms at maturity.

## 10.2 Embryology and Germination

Embryological morphology within *S. viridis* is typical of the Poaceae with embryos first being recognizable at the globular stage, where they appear as a dense, undifferentiated mass of cells embedded at the base of the developing endosperm (Fig. 10.1a). *Setaria viridis* also shows various synapomorphies unique to panicoid grasses that become more apparent at later developmental stages [Chap. 1]. Among these are the distinction between the sheathing tissue surrounding the embryonic root meristem (the coleorhiza) and the scutellum, which forms a haustorial attachment to the endosperm, resulting in a projection known as a scutellar tail (Fig. 10.1b). This distinction is absent in pooids and results in a “scutellar cleft” that is unique to panicoid species (Fig. 10.1b). Also characteristic of panicoid grasses is the

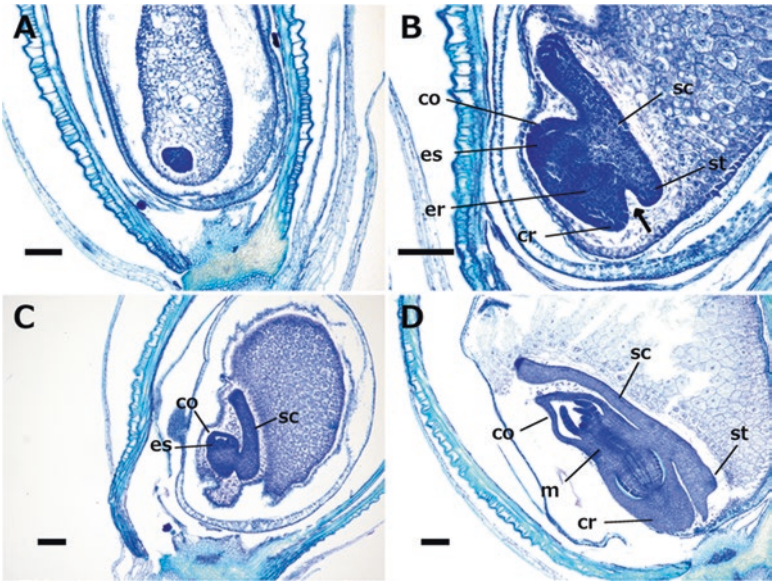


**Fig. 10.1** Embryological series of *S. viridis* stained with toluidine blue. (**a**) Globular stage embryo appearing as a dense, undifferentiated mass of cells embedded at the base of the developing endosperm. (**b**) Embryo during scutellum elongation and after polarity has been established given the presence of meristematic regions for the embryonic root and shoot. Note distinctive scutellar cleft visible between coleorhiza and scutellar tail (*arrow*). (**c**) Embryo after juvenile leaf primordia begin initiating from the embryonic shoot meristem axis. (**d**) Mature embryo prior to stratification, various panicoid features are recognizable such as the distinction between coleorhiza and scutellar tail as well as a mesocotylar internode. Coleoptile (co), coleorhiza (cr), embryonic shoot meristem (es), embryonic root meristem (er), mesocotylar internode (m), scutellum (sc), scutellar tail (st), scale bars at 100 μm

elongation of the mesocotyl, producing a prominent internode between the insertion of the scutellum and coleoptile along the embryonic axis (Fig. 10.1c, d) (Kellogg 2015). In keeping with general grass development, the embryonic apical meristem remains active for a prolonged period of time, resulting in the generation of at least two leaf primordia prior to maturation of the seed (Fig. 10.1c, d) (Kellogg 2000, 2015). In keeping with the Panicoideae disarticulation pattern, seeds are abscised below the insertion of glumes so that the entire spikelet axis acts as a diaspore (Doust et al. 2014). As a result of this mechanism, the caryopsis of *S. viridis* remains encased in the sclerified sterile lower lemma and the upper palea and lemma at germination.

Germination of *S. viridis* either on petri plates or within soil media occurs in a stereotyped manner within 2–3 days after imbibition (see below). If A10.1 seeds are planted immediately after harvest, germination rates are low and highly variable. However, following seed pretreatments, *S. viridis* A10.1 usually has high germination rates (>90%) across growing conditions. Pretreatment methods include a prolonged post-harvest storage period (>1 month) between senescence of the parent plant and sowing

of the offspring, a cold shock treatment (−80 °C for 2–3 days) (Mauro-Herrera et al. 2013), or addition of liquid smoke (Jose et al. 2014; Nelson et al. 2009). Growth conditions following germination are often crucial for a normal phenotypic response given A10's sensitivity to environmental conditions. As true for most $C_4$ grasses, *S. viridis* prefers high light and will often exhibit either stunted growth patterns or variable morphological patterns when grown under very low light levels (<200 μmol $m^{-2}$ $s^{-1}$).

As with other panicoids such as maize and sorghum, germination follows a stereotypical pattern in which the coleoptile and mesocotyl escape through the apex of the spikelet through the aperture created by the tips of the spikelet bracts. By contrast, early radicle development is more involved, given the presence of several sclerified bracts in *S. viridis*. In cases of successful germination, there is a highly predictable pattern of radicle emergence in which the coleorhiza punctures through a germination flap at the base of the upper lemma on the spikelet in order to exit and is shortly thereafter shed by the radicle (Fig. 10.2). This pattern has also been maintained in the domesticated cultivar *Setaria italica* (Keys 1949). Previous studies have indicated that *S. italica* germinates more quickly than *S. viridis*, but our own work with *S. viridis* (A10.1) and *S. italica* (B100 and Yugu1) accessions show the opposite, indicating that speed of germination is likely genotype dependent (Keys 1949). The prolonged functionality of the coleorhiza often allows it to remain discernible on the primary axis beyond germination (Fig. 10.2d), whereas, by contrast, the coleorhiza of maize does not elongate, and thus appears minute after germination (Hochholdinger et al. 2004).

*Early vegetative growth and juvenile to adult phasing when grown under standard conditions* (12:12 h day:night cycle, 28 °C day 22 °C night, humidity ~30 %, illumination ~30000 μmol $m^{-2}$ $s^{-1}$).

Following emergence from the soil, *S. viridis* plants sequentially produce four juvenilized leaves that lack notable intercalary meristem growth in their corresponding internodes. The exsertion of these first four leaves occurs rapidly, with the first leaf becoming fully expanded 2 days after emergence and the ligules of the second, third, and fourth leaves appearing 6, 8, and 10 days after emergence, respectively (Fig. 10.3). There is a general increase in size with each subsequent leaf, as these juvenile leaves form a transitional grade towards the mature leaves. Gross observations suggest a similarity between the juvenile phasing of leaves in Setaria and those of maize (Orkwiszewski and Poethig 2000; Sylvester et al. 2001). Mature leaves under high light conditions often appear to have larger leaf areas, due to increased blade width and length, a chaffy surface, and are often more rigid. Similar features are noted within mature maize leaves, where they correlate with increases in both trichome density and vascular development (Sylvester et al. 2001). The morphology of the juvenile leaves varies little when the plants are grown at varying light intensities and densities, suggesting that the underlying developmental process is largely insensitive to environmental variation. Blade morphology beyond the first four leaves shows some relationship with environmental factors, particularly light, as when adult glasshouse-grown *S. viridis* are compared to their growth room counterparts they often have wide blades and a chaffy surface whereas adult growth room leaves are only moderately differentiated from their juvenile counterparts.

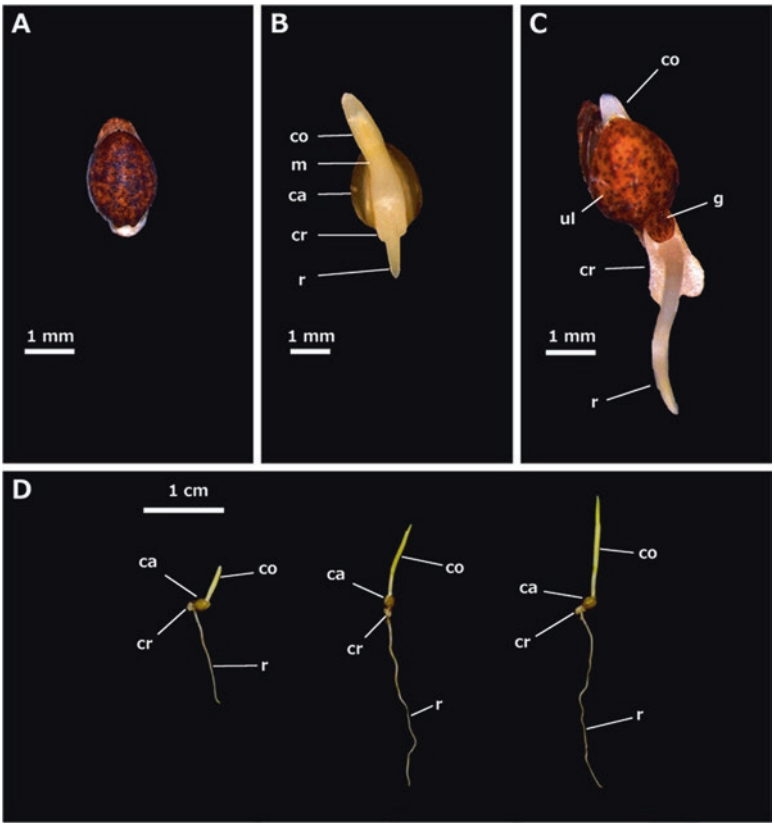


**Fig. 10.2** Early germination of *S. viridis* seeds. (**a**) Imbibed seed prior to germination displaying mottled patterning on the surface of the upper lemma. (**b**) Germinating seed which has been dissected to remove the upper lemma and palea, note small rectangular coleorhiza immediately adjacent to point of attachment of the radicle to the mesocotyl. (**c**) Germinating seedling where the germination flap facilitating the escape of the radicle can be seen at the base of the lemma along with the coleorhiza which has continued to elongate. (**d**) Seedlings several hours post-germination, growth of coleorhiza arrests shortly after first leaf exserts from coleoptile and vegetative growth begins. Caryopsis (ca), coleoptile (co), coleorhiza (cr), germination flap (g), mesocotyl (m), radicle (r), upper lemma (ul)

Under standard conditions, the shoot apical meristem of *S. viridis* appears to initiate new leaves roughly every ~2–3 days (Fig. 10.4). There appears to be no delay in bud outgrowth once buds have been initiated. Linear relationships between leaf initiation and exsertion have also been shown in sorghum and maize (Clerget et al. 2008; Abendroth et al. 2011). Except for abnormally fast exsertion rates of the first two leaves (which are products of embryogenesis) maize displays a continuous rate of leaf exsertion every ~5.5 days, a somewhat slower rate than *S. viridis* (Abendroth et al. 2011). It is hard to disentangle if these differences in continuous growth rates are specific to distinct lineages or simply a product of domestication, as the majority of work within panicoid development has largely been focused


**Fig. 10.3** Early vegetative growth of *S. viridis* under glasshouse conditions prior to axillary branch exsertion. Early vegetative growth at 2 days (**a**), 6 days (**b**), 10 days (**c**), 14 days (**d**), and 18 days (**f**) postemergence. Numbers in (**d**) and (**f**) indicate what serial position each leaf represents in the primary axis. Axillary buds begin to elongate within 2 weeks after emerging as seen upon dissection of the 14 day (**e**) and 18 day (**g**). Numbers listed below leaf blades and their corresponding axillary buds are the serial leaf positions denoted in (**d**) and (**f**). The leaf sheaths of the first leaf of day 14 (**d**) and the third leaf of day 18 (**f**) have been pushed away from the stem by the outgrowth of the first tiller (**e**) and third tiller (**g**), respectively

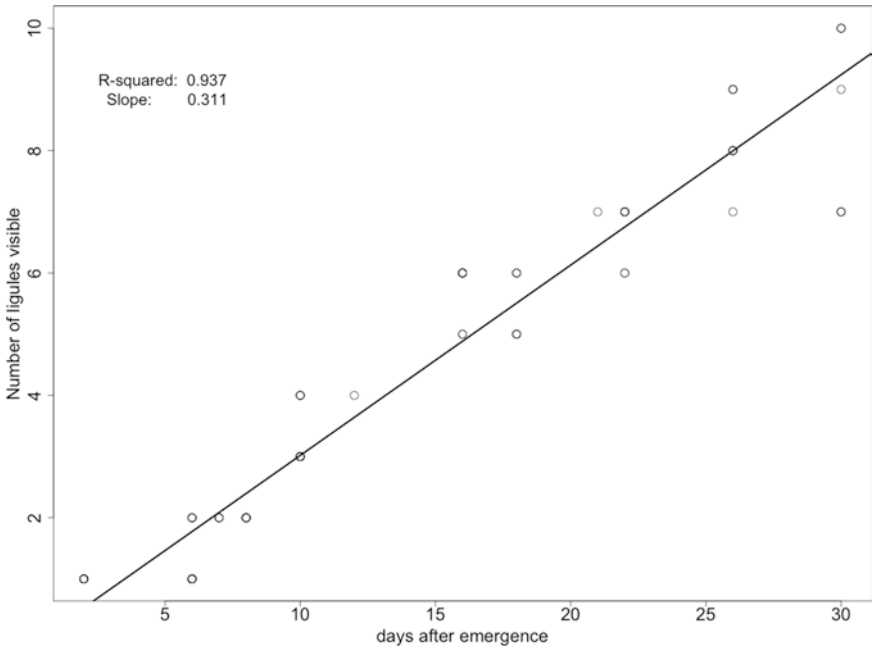


**Fig. 10.4** Graph displaying periodic plastochron of *S. viridis* in which rate of leaf exsertion is compared to absolute time. $m = 0.311$, $R^2 = 0.9362$

within domesticated cereals. However, observations of *S. italica* accessions B100 and Yugu1 show rates of leaf exsertion that are more comparable to those of *Z. mays,* suggesting a potential relationship between rate of organ initiation/elongation and domestication.

## 10.3 Axillary Branching and Root Architecture

Axillary branching occurs in two distinct phases of growth in *S. viridis,* with tillers being initiated from the basal nodes (mostly bearing juvenile leaves) and aerial branches being initiated from the nodes separated by elongated internodes along the culm and tillers. Aerial branching is often the most pronounced after the main culm has transitioned to flowering (Doust et al. 2004). The basal growth phase produces several axes (tillers) that have the potential to develop a lateral root system independent of the primary axis. Often only the first four axillary axes are near enough to the soil surface to successfully establish an independent root system and, of these, the internodes immediately associated with the first two axes are the primary sites for adventitious root initiation. Under both glasshouse conditions (14 h days with >1000 μmol $m^{-2}$ $s^{-1}$ light at peak hours) and growth room conditions (12 h days and constant light ~300 μmol $m^{-2}$ $s^{-1}$), the first discernible tillers appear in the axils of

leaves 1 and 2 at roughly 12–14 days postemergence (Fig. 10.3). Elongation of axillary buds is first recognizable by the surrounding leaf sheath being dislodged from the stalk by the rapid outgrowth of the bud (Fig. 10.3d–g). Despite being the first axes to elongate, the first and second tillers are only able to grow for a few days, producing a few juvenile leaves, before ceasing growth entirely (Fig. 10.5). The first tiller appears to always follow this pattern but the second tiller can sometimes continue growth (Fig. 10.5). This pattern becomes even more pronounced as the third and fourth tillers begin to elongate as well, often quickly overtaking their basal counterparts (Figs. 10.3 and 10.5). There are also developmental differences in the types of leaves being initiated on the first and second tillers, which primarily bear juvenilized leaves, compared to that of the third and fourth, where the first leaf is often juvenilized, and followed thereafter by mature leaves, suggesting potential differences in how these axes are developmentally canalized (Fig. 10.5). Moreover, when compared to the first two tillers, the rates of growth and leaf emergence are more uniform for the higher axes.

In accordance with studies in other grasses, the establishment of the postembryonic root system often shows a positive correlation with tillering in *S. viridis* (Manske and Vlek 2002). Crown roots often begin to appear within a week of axillary meristem elongation, so that the crown root system occupies a comparable volume to the shoot system at flowering (Fig. 10.5). The intensity of crown root initiation varies between phytomers, often with circumscissile rings of roots initiating from the first and second internodes while only sparse root initiation is visible in the third and fourth. The exsertion of these roots can have notable effects on the shoot system, with the leaf sheaths of leaves 1 and 2 often becoming severely damaged and rendered nonfunctional, and their corresponding tillers forced out of a distichous arrangement (Fig. 10.6). In some cases, the damage is more severe, with root elongation inflicting structural damage to the first and second tillers themselves. In contrast, the scarce root outgrowth from the third and fourth internodes limits damage to the surrounding sheaths and thus allows the blades of leaves 3 and 4 to remain functional until senescence. *S. viridis* shoot growth is sensitive to constrictions placed on root system growth, and plants may flower earlier if root volumes are restricted (Table 10.1).

## 10.4 Transition to Flowering and Inflorescence Morphology

The transition of the vegetative shoot apical meristem into the reproductive fate of a inflorescence meristem occurs when the first mature leaves are visible externally (or shortly after). Previous work in maize has suggested that a set number of phytomers is required for reproductive competency to be reached, enabling the transition of the shoot apical meristem to the inflorescence meristem (Irish and Jegla 1997). In a manner similar to many other grasses, floral transition is often first discernible in Setaria as the elongation of the meristematic axis into a pin-like structure, which causes it to exsert beyond the shelter of the leaf primordia that cover it

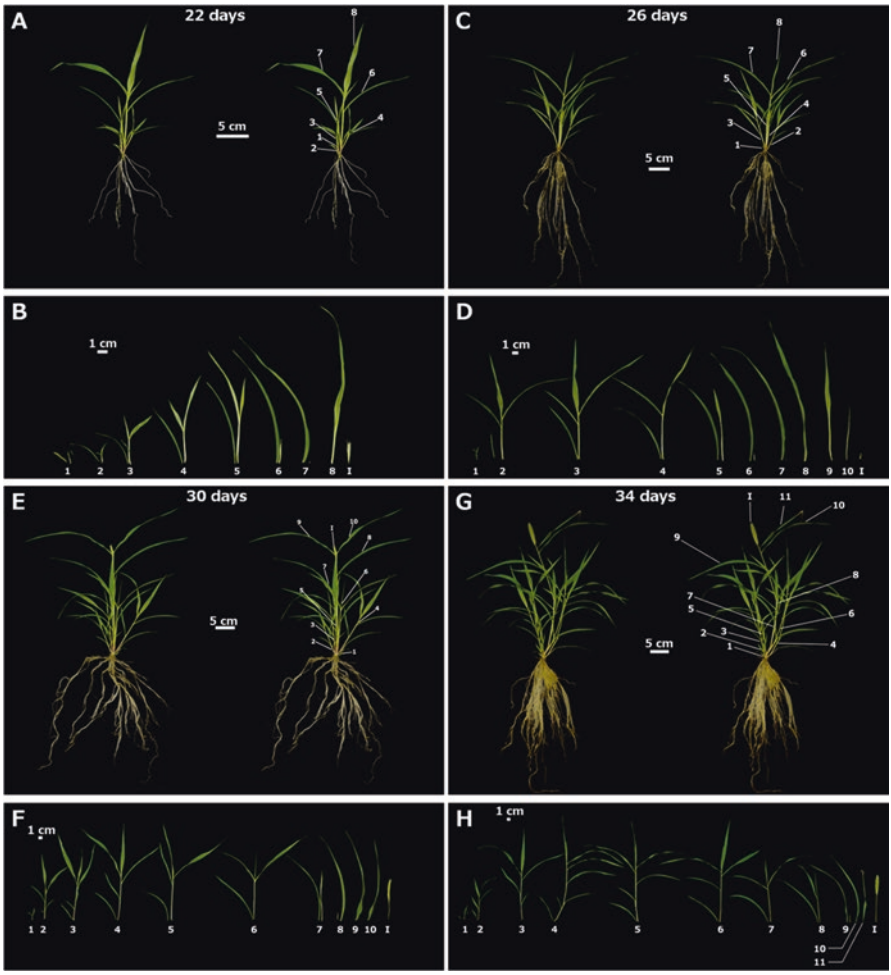


**Fig. 10.5** Later vegetative growth of *S. viridis* under glasshouse conditions showing patterns of axillary branching up until flowering at 22 days (**a**), 26 days (**c**), 30 days (**e**), and 34 days (**g**) postemergence. Numbers in (**a**), (**c**), (**e**), and (**g**) indicate what serial position each leaf represents in the primary axis. Upon dissection of axillary branches, the increased growth effort in the axillary branches of the leaf 3 and 4 becomes apparent with the branches from the axils of leaves 1 and 2 showing little growth by comparison over developmental time (**b**), (**d**), (**f**), and (**h**). As the axillary branches of higher axes above the fourth leaf begin to elongate, the axillary branches usually form a developmental grade, recapitulating the order in which they were produced from the shoot apical meristem (**d**), (**f**), and (**h**). Numbers listed below leaf blades and their corresponding axillary branches in (**b**), (**d**), (**f**), and (**h**) are the serial leaf positions denoted in (**a**), (**c**), (**e**), and (**g**), respectively

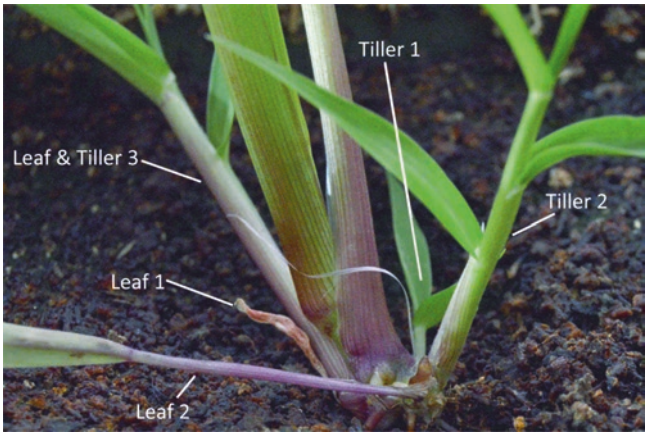


**Fig. 10.6** Structural changes in first and second phytomers resulting from crown root exsertion. Severe structural damage is present to sheath of the second leaf, rendering it nonfunctional. Note differences in orientation between leaves 1 and 2 compared to tillers 1 and 2, showing severity of displacement of the tillers by the secondary root system

**Table 10.1** Variation in trait values for *S. viridis* in varying growth conditions, compared to standard conditions

| Variable | Values | Time to flowering | Height | Branch number | Leaf number |
|---|---|---|---|---|---|
| Plant density (low to high) | Field (900 $cm^2$, 55 $cm^2$) | Same | Increased | Decreased | Not measured |
| Plant density (low to high) | Greenhouse (55 $cm^2$, 30 $cm^2$) | Same | Increased | Decreased | Not measured |
| Photoperiod (short to long) | Growth chamber (8, 12, 16 h light) | Increased | Increased | Increased | Increased |
| Root volume (small to large) | Greenhouse (1400 µmol $m^{-2}$ $s^{-1}$: 115, 230, 345 $cm^3$) | Increased | Increased | Same | Same |
| Root volume (small to large) | Greenhouse (400 µmol $m^{-2}$ $s^{-1}$: 115, 230, 345 $cm^3$) | Increased | Increased | Decreased | Same |
| Light intensity (high to low) | Greenhouse (1400, 400 µmol $m^{-2}$ $s^{-1}$) | Increased | Increased | Increased | Same |
| Light intensity (high to low) | Growth chamber (250, 115 µmol $m^{-2}$ $s^{-1}$) | Increased | Same | Decreased | Increased |

These data come from unpublished growth trials (Doust, unpublished). Only trends for environmental variables that were varied within trials are noted, along with the values for those variables

during its vegetative phase (Fig. 10.7a, b.). Early inflorescence development in Setaria bears a striking resemblance to the pistillate ears of maize at a developmentally analogous stage with primary branches being initiated acropetally (Fig. 10.7c). Unlike the paired spikelet meristems of maize resulting from these primary branches, Setaria undergoes a brief phase of distichous fractal-like branching along each primary branch resulting in a complex but highly repetitive morphology (Fig. 10.7d, e, h) (Doust and Kellogg 2002). The meristems produced from this process are then able to transition into either the sterile fate of bristles or fertile fate of spikelets (Fig. 10.7f). Bristles appear to be patterned from aborted spikelet meristems and are often first recognizable based on their elongated pedicels (when compared to fertile spikelets) and the circumcissile scar that forms around the base of the meristematic dome (Fig. 10.7f). Shortly after the formation of this scar, the spikelet meristem of the bristle often becomes necrotic and collapses, shedding shortly thereafter so that the apex of each bristle contains only a blunt stump, where the meristematic dome was previously attached (data not shown). By contrast, the spikelet meristems often have a short pedicel and develop as typical two-flowered panicoid spikelets, with the lower floret usually sterile (Fig. 10.7f, g). There has been some suggestion that the branching patterns of Setaria are an elaborated version of the paired spikelet branching pattern found within the Andropogoneae (Fig. 10.7h) (Zanotti et al. 2010). This results in the primary axis for each short branch order aborting into a sterile pedicellate axis while its recently generated axillary axis is retained as a fertile sessile spikelet (Fig. 10.7h). The combination of these growth patterns results in the characteristic arrangement of bristles surrounding the fertile spikelets of Setaria at maturity (Fig. 10.7i) (Doust and Kellogg 2002; Doust et al. 2005).

## 10.5 Environmental Sensitivity

Perhaps the most striking feature of *S. viridis* compared to other panicoid models such as maize and Sorghum is that it is a wild species and thus retains much of its phenotypic plasticity. There are various environmental stimuli which have been noted to cause phenotypic shifts, such as soil volume and depth, day length, and light quality (Doust, unpublished), as well as water stress (Fahlgren et al. 2015) (Chap. 16). Moreover, many of these developmental decisions are established early in the *Setaria* life cycle.

As an indication of the range of phenotypes that can be expected under varying growth conditions, we have compiled data from multiple unpublished trials for architectural and flowering time traits in *S. viridis* (Table 10.1). These trials vary widely in light intensity, root volume, and plant density, as well as photoperiod. Some of the differences in phenotype can be attributed to a shade avoidance response, as in the field and greenhouse density trials, where height increases, branching decreases, but flowering time stays the same. Other environmental changes affect flowering time as well, such as changing photoperiod, root volume, or light intensity. Interestingly, phenotypic responses can vary depending on the levels of an environmental variable that are

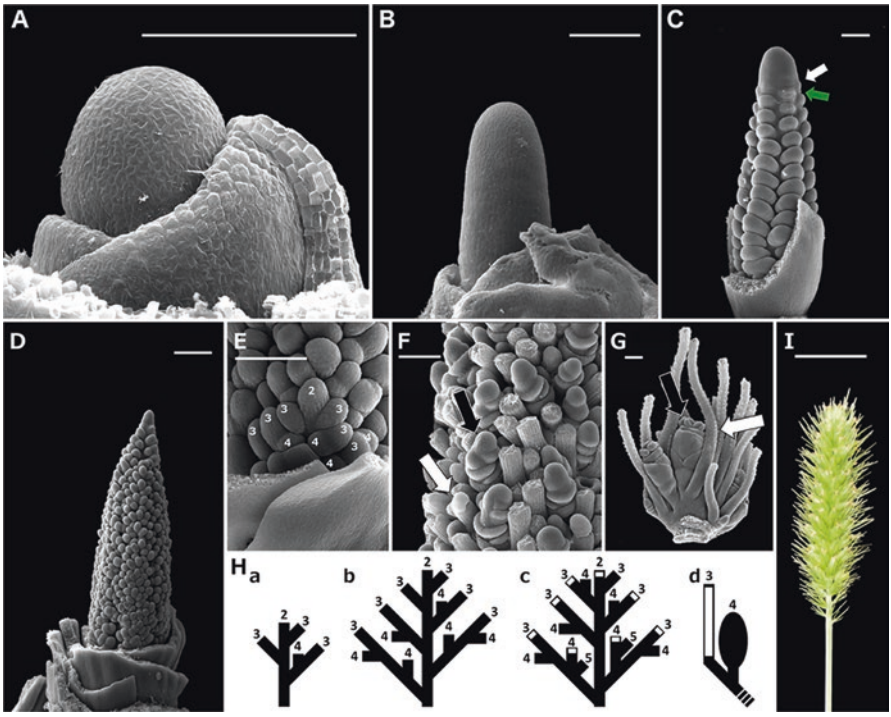


**Fig. 10.7** Floral transition of the apical meristem of *S. viridis*. (**A**) Vegetative meristematic dome with sheathing leaf primordia removed. (**B**) Inflorescence meristem shortly after its floral transition causing a shift in morphology from a small dome shielded by primordia to a pin-like structure that elongates beyond the younger primordia which form a characteristic hood over the vegetative axis. (**C**) Primary branching phase of inflorescence development, note the small scale-like bracts (*green*) subtending each branch meristem (*white*) shortly before they elongate. (**D**) Formation of distichous secondary branch primordia on the primary branch axes. (**E**) A single axillary branch annotated so that each branch order is numbered based on its rank to the primary axis. (**F**) The branch meristem to spikelet meristem transition follows a typical panicoid pattern in which fertile axes (*black*) are generated that bear one or two floret meristems, or alternatively, the meristematic axis is aborted and shed so that the barren axis that eventually develops into a bristle (*white*). (**G**) An isolated axillary branch at a later stage of floral development showing a spikelet pair consisting of a fertile sessile spikelet (*black*) and a sterile pedicellate axis (*white*). (**H**) Illustration outlining branching patterns and spikelet to bristle transitions within Setaria. (**a**) Early stage of axillary branch elongation in which third-order axes are generated acropetaly along the second-order axis. (**b**) Later developmental stage in which the patterns of the fourth branch order is recognizable and is comparable to the architecture visible in (**E**). (**c**) Transition of axes to a spikelet fate occurs shortly after (**b**) with terminal axes (*white*) differentiating into a sterile bristle fate whereas axillary axes are retained as fertile spikelet fates. (**c**) This illustration interprets the spikelet and bristle fates as analogous to a paired spikelet morphology in maize. (**d**) A representation of a paired spikelet to compare to (**G**) in which the terminal axis (*white*) has developed into a bristle, whereas the axillary axis is retained as a spikelet (*black*) and the dashed line represents the point of attachment to the lower order axis. (**I**) Panicle of *S. viridis* at maturity bearing both fertile spikelets and bristles. Scale bars for A–G = 100 μm, I = 1 cm

applied. This appears to be the case for light intensity, where variation in physiologically meaningful levels of variation (1400, 400 μmol $m^{-2}$ $s^{-1}$), increases flowering time, height, and branch number, but where much lower levels of light (250, 115 μmol $m^{-2}$ $s^{-1}$) increase flowering time, do not affect height, and decrease branching. The sensitivity of *S. viridis* to environmental variation has the potential to be a useful tool in understanding plant–environment responses. We have made some progress in understanding the effect of photoperiod on growth and development in Setaria [(Orkwiszewski and Poethig 2000), Chap. 12], but much work still needs to be done to unravel the effects of carbon gain and light amount on architecture and flowering time.

## 10.6 Discussion

In this chapter, we have set out developmental growth patterns for *S. viridis* A10.1 from germination to flowering, with most emphasis on changes in plant architecture during growth. Like maize, Setaria undergoes a distinct juvenile phase in which leaves appear to be smaller at lower nodes and their corresponding internodes also reduced in size. The axillary buds of these leaves are often the most likely to develop into tillers given their proximity to the soil. Also like other cereal models, leaf initiation and exsertion are highly periodic, once a meristematic axis has committed to organogenesis. There is also evidence that not all vegetative axes are equal, with the first and occasionally the second tillers arresting growth shortly after starting to elongate. We also summarize the results of several growth trials between various experiments in order to emphasize the plastic nature of this accession. It is important to be aware of the labile nature of vegetative growth in Setaria, as varying experimental conditions will change phenotypic outcomes, irrespective of treatments.

The other source of variation explored in this chapter is developmental, and the stages described above provide insight into growth patterns in *S. viridis* A10.1. The short life cycle means that the ontogenetic program of an individual is determined soon after germination, with developmental decisions related to branching and flowering occurring as early as the first 2 weeks following emergence. Early inflorescence development bears a superficial similarity to the ears of maize although the fractal-like branching patterns of the axillary inflorescence branches and the conversion of spikelet meristems into bristles through abortion of their meristems are both characteristic of Setaria. The circumcissile scars of *S. viridis* bristles resulting in spikelet meristem abortion have a resemblance to the abscission zones of spikelets although such an interpretation is at present pure speculation (Hodge and Kellogg 2016).

The underlying genetics driving plant architecture have been explored in several studies. These have implicated various architectural regulators such as *MONOCULM1*, *SEMI-DWARF1*, and *teosinte branched1* in phenotypic variation that was selected upon during the domestication of *S. italica* [(Manske and Vlek 2002), (Mauro-Herrera and Doust 2016), Chap. 12]. Much remains to be done to understand the relationship between developmental timing, environmental sensitivity, and perception, and *S. viridis* promises to be a rich source of variation and insights into these questions.

**Acknowledgements** We would like to thank E. A. Kellogg for permitting us to use the histological data of *S. viridis* embryology in this chapter, as well as M. S. Box who generated a scanning electron micrograph used in this chapter (Fig. 10.7g). We would also like to thank Lisa Whitworth at the Oklahoma State University Microscopy Core Facility for help with scanning electron microscopy and finally, Michael Malahy, Rene Mitchell, Josh Keegan, Cara Stephens, and Emily Harris who generated unpublished data on Setaria under various growth conditions.

## References

Abbo S, van-Oss RP, Gopher A, Saranga Y, Ofner I, Peleg Z. Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes. Trends Plant Sci. 2014;19(6):351–60.

Abendroth L, Elmore RW, Boyer MJ, Marlay SK. Corn growth and development. Ames: Iowa State University; 2011.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X, Kellogg EA, Van Eck J. *Setaria viridis*: a model for C4 photosynthesis. Plant Cell. 2010;22:2537–44.

Clerget B, Dingkuhn M, Goze E, Rattunde HFW, Ney B. Variability of phyllochron, plastochron and rate of increase in height in photoperiod-sensitive *Sorghum* varieties. Ann Bot. 2008;101:579–94.

Doust AN, Kellogg EA. Inflorescence diversification in the panicoid "bristle grass" clade (Paniceae, Poaceae): evidence from molecular phylogenies and developmental morphology. Am J Bot. 2002;89(8):1203–22.

Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA. Genetic control of branching in foxtail millet. Proc Natl Acad Sci U S A. 2004;101(24):9045–50.

Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA. The genetic basis for inflorescence variation between foxtail and green millet (Poaceae). Genetics. 2005;169:1659–72.

Doust AN, Mauro-Herrera M, Francis AD, Shand LC. Morphological diversity and genetic regulation of inflorescence abscission zones in grasses. Am J Bot. 2014;101(10):1759–69.

Fahlgren N, Feldman M, Gehan MA, Wilson MS, Shyu C, Bryant DW, Hill ST, McEntee CJ, Warnasooriya SN, Kumar I, Ficor T, Turnipseed S, Gilbert KB, Brutnell TP, Carrington JC, Mockler TC, Baxter I. A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in *Setaria*. Mol Plant. 2015;8(10):1520–35.

Hochholdinger F, Woll K, Sauer M, Dembinsky D. Genetic dissection of root formation in maize (*Zea mays*) reveals root-type specific developmental programmes. Ann Bot. 2004;93:359–68.

Hodge JG, Kellogg EA. Abscission zone development in *Setaria viridis* and its domesticated relative, *Setaria italica*. Am J Bot. 2016;103(6):1–8.

Irish EE, Jegla D. Regulation of extent of vegetative development of the maize shoot meristem. Plant J. 1997;11(1):63–71.

Ishida Y, Hiei Y, Komari T. Agrobacterium-mediated transformation of maize. Nat Protoc. 2007;2(7):1614–21.

Jose S, Wong MK, Tang E, Dinneny JR. Methods to promote germination of dormant *Setaria viridis* seeds. PLoS One. 2014;9(4):e95109.

Kellogg EA. The grasses: a case study in macroevolution. Annu Rev Ecol Syst. 2000;31:217–38.

Kellogg EA. Evolutionary history of the grasses. Plant Physiol. 2001;125:1198–205.

Kellogg EA. Flowering plants monocots—Poaceae. In: Kubitzki K, editor. New York: Springer; 2015.

Keys CE. Observations on the seed and germination of *Setaria italica* (L.) Beauv. Trans Kans Acad Sci. 1949;52(4):474–7.

Manske GGB, Vlek PLG. Root architecture—wheat as a model plant. In: Waisel Y, Eshel E, Beeckman T, Kafkafi U, editors. Plant roots: the hidden half. 3rd ed. New York: Marcel Dekker; 2002. p. 382.

Martins PK, Nakayama TJ, Ribeiro AP, Dias Brito da Cunha BA, Nepomuceno AL, Harmon FG, Kobayashi AK, Molinari HBC. *Setaria viridis* floral-dip: a simple and rapid Agrobacterium-mediated transformation method. Biotechnol Rep. 2015;6:61–3.

Mauro-Herrera M, Doust AN. Development and genetic control of plant architecture and biomass in the Panicoid grass, *Setaria*. PLoS One. 2016;11(3):e0151346.

Mauro-Herrera M, Wang X, Barbier H, Brutnell TP, Devos KM, Doust AN. Genetic control and comparative genomic analysis of flowering time in Setaria (Poaceae). G3. 2013;3(2):283–95.

Nelson DC, Riseborough J, Flematti GR, Stevens J, Ghisalberti EL, Dixon KW, Smith SM. Karrikins discovered in smoke trigger *Arabidopsis* seed germination by a mechanism requiring gibberellic acid synthesis and light. Plant Physiol. 2009;149:863–73.

Orkwiszewski JAJ, Poethig RS. Phase identity of the maize leaf is determined after leaf initiation. Proc Natl Acad Sci U S A. 2000;97(19):10631–6.

Sylvester AW, Parker-Clark V, Murray GA. Leaf shape and anatomy as indicators of phase change in the grasses: comparison of maize, rice, and bluegrass. Am J Bot. 2001;88(12):2157–67.

Van Eck J, Swartwood K. Setaria viridis. In: Wang K, editor. Agrobacterium protocols. New York: Springer; 2014.

Zanotti CA, Pozner R, Morrone O. Understanding spikelet orientation in Paniceae (Poaceae). Am J Bot. 2010;97(5):717–29.

# Chapter 11
# *Setaria viridis*: A Model for Understanding Panicoid Grass Root Systems

**Jose Sebastian and José R. Dinneny**

**Abstract** Roots are essential for plant survival on land. Understanding how root traits relate to overall crop yield will be key to sustainably supporting an ever-expanding population. For global food and biomass production, members of the grass family (Poaceae) contribute the lion's share; however, our understanding of grass root biology remains rather poor. Among grasses, Panicoideae subfamily grasses are among the most agronomically important groups of plants. Recently, *Setaria viridis* (Setaria) has emerged as a new genetic model system for Panicoideae grasses. Setaria characteristics such as a relatively small genome, fast life cycle, ease of growing under controlled conditions, and remarkable drought tolerance make it an excellent plant model system to study various aspects of grass biology. Setaria has a typical grass root system architecture composed of a primary, crown and lateral roots, making it feasible to conduct systematic analyses that elucidate general physiological mechanisms of broad relevance. In this chapter, we give an overview of root systems in grasses and Panicoideae grasses and provide a detailed description of the *Setaria viridis* root system highlighting different root types and their internal cellular organization.

**Keywords** Monocot root systems • Root system biology • Panicoid grasses • *Setaria viridis*

## 11.1 Introduction

### *11.1.1 The Importance of Roots*

Roots are a multifunctional organ system through which plants obtain most of their water and nutrients required for growth and development. Being sessile organisms, plants depend heavily on their root systems and their adaptive responses to

J. Sebastian (✉) • J.R. Dinneny
Department of Plant Biology, Carnegie Institution for Science,
260 Panama Street, Stanford, CA 94305, USA
e-mail: jsebastian@carnegiescience.edu; jdinneny@carnegiescience.edu



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_11

optimally extract soil reserves. Important adaptive responses to exogenous stimuli include alterations in root growth rate and branching patterns. This is particularly important when soil resources are either depleted or inaccessible due to a multitude of stresses (Lynch 1995; Grossman and Rice 2012; Postma et al. 2014a). Besides nutrient uptake, roots are also important in anchorage, storage of nutrient reserves, and forming mutually beneficial microbial associations that are often critical for proper plant growth (discussed in Chap. 14, and reviewed by (Bulgarelli et al. 2013)). Considering the rapid increase in the human population and associated nutritional demands, it is of paramount importance to improve our agriculture production through sustainable means. Crop plants with better root traits are predicted to be the next main source of improvement in our agriculture (Hammer et al. 2009; White et al. 2013; Rogers and Benfey 2015). Progress has been made in understanding the biology of roots and root systems; however, most of this is knowledge pertaining to eudicot models, particularly *Arabidopsis thaliana* (reviewed by (Petricka et al. 2012)). While commonalities exist between monocot and eudicot root systems, important differences exist, which require investigation. Knowledge of the genetic networks that drive root growth and response to stress in monocot plants will ultimately help in generating crop plants with superior root systems that are better equipped to grow under resource-limiting agricultural conditions (Kong et al. 2014).

### *11.1.2 Significance of Panicoideae Grasses*

Agronomically, members of the Panicoideae subfamily of grasses are particularly important. There are approximately 3300 species belonging to this subfamily (Grass Phylogeny Working Group (GPWG 2001)) with a truly global presence (Giussani et al. 2001). Many important feed, fodder, and fuel crops such as maize (*Zea mays*), sorghum (*Sorghum bicolor*), foxtail millet (*Setaria italica*), common millet (*Panicum miliaceum*), pearl millet (*Pennisetum glaucum*), sugarcane (*Saccharum officinarum*), Miscanthus (*Miscanthus giganteus*), and switchgrass (*Panicum virgatum*) are all members of this clade. Many of these plants exhibit traits that will be vital in a world facing critical challenges on several fronts including global warming, water scarcity, and degradation of arable lands. For example, crops such as sorghum and millet are capable of growing in arid and semiarid environmental conditions often found in the most economically challenged parts of the globe (Pray and Nagarajan 2002). Miscanthus and switchgrass are fast growing plants capable of cultivation on marginal lands and may assist in supporting the ever-increasing energy demands of the world through biofuel production (Lewandowski et al. 2003; Khanna et al. 2008). As an added advantage, many of the agronomically important Panicoideae grasses have evolved a $C_4$ photosynthesis system, which is more efficient at converting solar energy to biomass owing to their ability to concentrate $CO_2$ and reduce photorespiration, thus improving agricultural productivity (Giussani et al. 2001; Zhu et al. 2008; Osborne et al. 2014).

### *11.1.3 Comparisons Between Eudicot and Grass Root Systems*

There are distinct differences between eudicot and grass root system both in terms of root system structure and internal cellular organization. One of the key differences between these two root systems is the presence of additional root types in monocots such as the crown/nodal roots (Fig. 11.1a–c; Table 11.1). These shoot-borne roots together with associated lateral roots form the bulk (90–95 %) of the adult plant root system commonly referred to as a fibrous root system (Metcalfe and Nelson 1985; Hoppe et al. 1986; Hochholdinger et al. 2004a). In contrast, shoot-borne roots are

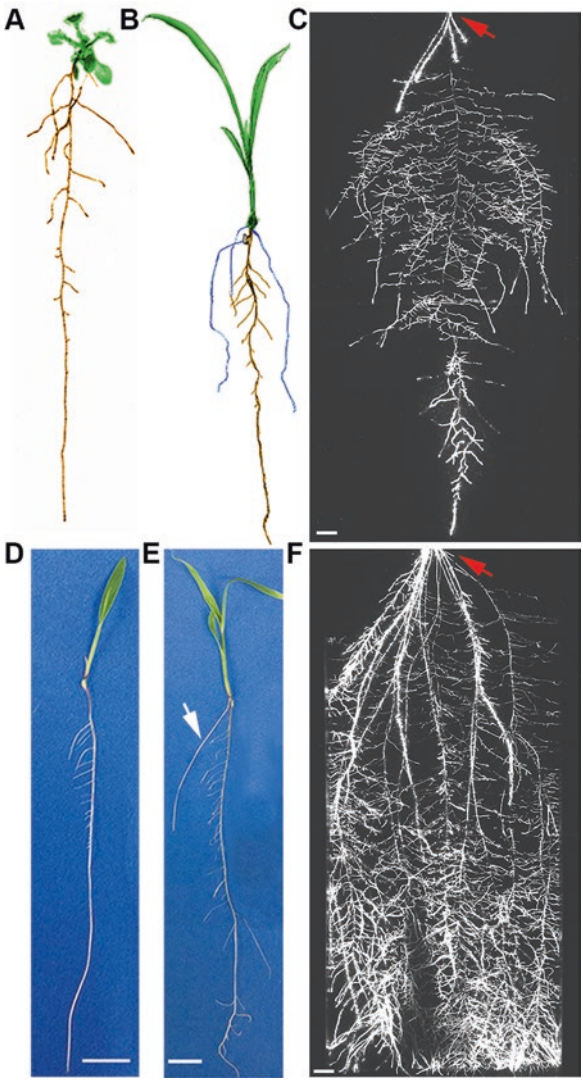


**Fig. 11.1** *Setaria viridis* root system. (**a**, **b**) A diagrammatic representation of a eudicot (Arabidopsis) root system and a monocot (Setaria) root system. Embryonic root system is colored in *yellow* and *blue* represents shoot-borne postembryonic root system. (**c**) Luminescence-based image of *S. viridis* root system at 17 days after sowing (DAS). *Arrowhead* indicates the crown region. (**d**) Setaria root system at 4 DAP. Seedlings germinated on agar media. (**e**) Setaria root system at 7 DAP showing the emergence of a crown root from the coleoptilar node. *Arrow head* indicates crown root. (**f**) Luminescence-based image of *S. viridis* root system at 29 DAS. *Arrowhead* indicates the crown region. Scale bars: 1 cm

**Table 11.1** Common root types in panicoid grass root systems

| Root type | Site of emergence | Time of primordia initiation | Root characteristics |
|---|---|---|---|
| Primary root | Basal pole of embryo | Embryogenesis | Continuation of the radicle. First root to emerge following seed germination |
| Seminal roots | Scutellar node | Embryogenesis | Emerge few days post seed germination. Exact number per seedling varies. Not every grass species form |
| Crown roots | Shoot nodes (underground) | Post-embryogenesis | Shoot-borne roots form a major constituent of postembryonic root system in grasses. Exhibit heteroblasty |
| Brace roots | Shoot nodes (aboveground) | Post-embryogenesis | Shoot-borne roots. Important in lodging resistance. Common in maize |
| Sett roots | Shoot nodes (underground) | Post-embryogenesis | Highly branched, thin roots found in sugarcane. Crucial in the early establishment of the plant (sett) |
| Shoot roots | Shoot nodes (underground) | Post-embryogenesis | Thick, less branched roots. Common in sugarcane where they gradually replace sett roots |
| Buttress roots | Shoot nodes (aboveground) | Post-embryogenesis | Thick roots common in Sorghum. Important in providing lodging resistance |
| Superficial roots | Shoot nodes (underground) | Post-embryogenesis | Highly branched, thin roots. Common in sugarcane. Usually, grow laterally near the soil surface |
| Rope roots | Shoot nodes (underground) | Post-embryogenesis | Agglomerations of shoot roots grow vertically into deep soil. Important in nutrient acquisition and anchorage |
| Lateral roots | Pericycle and endodermis cells of all roots | Post-embryogenesis | Roots formed from other roots. Exhibits a high degree of branching. Important in water and nutrients uptake |
| Adventitious roots | Mesocotyl | Post-embryogenesis | Usually, not part of the normal root development program. Emerge commonly as a response to external stimuli |

typically not a major root type in eudicot root systems, where the primary root and associated several orders of lateral roots constitute the majority of the root system (referred to as a taproot system). The taproot system is adapted to grow deeply into the soil, while the fibrous root system is usually more shallow rooted and suited to exhaustive utilization of water and nutrients in the uppermost levels of the soil. There are also distinct differences in root anatomy between individual grass and dicot roots. There is usually no vascular cambium (a type of lateral meristem) developed in the grass root stele, and therefore these roots exhibit no secondary growth, unlike many eudicot roots (Hochholdinger et al. 2004b; Scarpella and Meijer 2004). There are also differences in xylem vessels with regard to shape and number. The origin of lateral roots also differs; in eudicots they are formed from pericycle cells in all root types, whereas in grasses, cell divisions in both the pericycle and endodermis contribute to

lateral root formation (Hochholdinger and Zimmermann 2009; Orman-Ligeza et al. 2013). At a larger scale of organization, however, there are similarities between grass and eudicot roots. In both groups, the root apical meristem, which harbors a stem cell niche at the apex sustains continued root growth by the iterative processes of cell division, followed by cell elongation and differentiation (Bennett and Scheres 2010; Sebastian and Lee 2013). There are two pools of stem cells present at the stem cell niche, the proximal stem cells or vascular initials, (progenitors of vascular tissue) and the distal stem cells or the collumella initials (progenitors of root cap tissue). These stem cells encompass a mitotically inactive group of cells referred to as the quiescent center (QC). Longitudinally, based on cell length, each grass or eudicot root type can be subdivided, starting from the root apex, into three sequentially arranged developmental zones: meristematic, elongation, and differentiation/maturation.

Crown roots emerge from the shoot nodes in a sequential manner beginning from the lower nodes at the shoot base (Onderdonk and Ketcheson 1972). These nodes are collectively referred to as the seedling crown. Although the overall structure and function of different root types in the root system are similar, there is evidence suggesting differences in their biology and responses to external cues (reviewed by (Bellini et al. 2014)). Many studies have reported that the primary, crown, and lateral roots often respond differently to soil/environmental conditions. For example, phosphorus deficiency in rice promotes primary and crown root elongation, while it suppresses lateral root growth (Zhou et al. 2008; Hu et al. 2011; Dai et al. 2012). This is in contrast to most other plants where phosphorus deficiency stimulates both lateral and adventitious root growth while suppressing primary root elongation (Bellini et al. 2014). Similarly, adventitious and lateral root growth is found to be more sensitive to the stress hormone abscisic acid (ABA), while the primary root is less sensitive to ABA (Bellini et al. 2014; De Smet et al. 2003; Duan et al. 2013). Moreover, there are mutants isolated in maize and other species that affect only a particular root type, further highlighting their underlying differences (Hochholdinger et al. 2004b; Bellini et al. 2014).

As mentioned above, most of our understanding of root biology is from Arabidopsis. However, Arabidopsis does not develop an extensive shoot-borne root system as grasses do. Thus, though useful, it is difficult to fully extend knowledge from Arabidopsis to grass species. Therefore, to elucidate grass root biology, it is imperative to develop model systems in the grasses.

## 11.2 Root Systems in Panicoideae Grasses

The high rate of productivity observed in grasses is, in part, due to their elaborate root systems, which are well adapted for exploring and extracting soil resources and water (Lynch 1995; Aiken and Smucker 1996). Therefore, improvements in our ability to grow panicoid grasses in agriculture will likely require a better understanding of their root systems; but which model will provide the greatest insight? Over the years, considerable efforts have been made to develop rice as a genetic model system, and this has contributed significantly to our understanding of grass root biology

(Itoh et al. 2005; Coudert et al. 2010). However, although the overall root system architecture is similar between rice and panicoid grasses, there is a limit as to how much knowledge we can transfer between the two. Rice is adapted to growth in flooded paddies where hypoxia is an important stress whereas the panicoid grasses are mostly habituated to arid/semiarid soil conditions where water deficit is more likely to be encountered. Moreover, there may be variations in root system growth dynamics due to the differences in $C_3$ (rice) versus $C_4$ (many panicoid grasses) photosynthesis. Thus, it is logical to explore the root system of panicoid grasses to get a clearer picture of their growth dynamics, stress responses, and underlying genetic regulatory networks. We will next provide a comparative description of root system characteristics in four different panicoid grasses. Although there is an underlying commonality that involves similarity in root system structure, function, and development, there are distinct root system features unique to each of these grasses. Most of the features illustrated below are species-specific root system adaptations acquired in order to optimize growth and survival in different ecological niches.

### *11.2.1 Root System Development in Maize*

Most of our understanding of root system structure and development in panicoid grasses is from studies of maize. The maize root system can be broadly categorized as embryonic or postembryonic based on the time of root emergence (Hochholdinger et al. 2004b; Feldman 1994; Hochholdinger 2009). Embryonic roots are those that form during embryogenesis through seed germination phases and include a single primary root and seminal roots. Unlike in eudicots, the primary root is formed endogenously from the basal pole of the embryo and thus must penetrate through the embryonic tissues before its emergence near the tip end of the kernel (Hochholdinger et al. 2004b; Feldman 1994). This internal origin of the primary root is a characteristic feature of monocotyledonous plants belonging to the true grass family Poaceae/Gramineae (Hochholdinger 2009). However, the significance of this internal origin is not obvious compared to the site of origin for the eudicot primary root, which is patterned from the outer tissue layers of the embryo (Bennett and Scheres 2010). A varying number of seminal roots are a distinguishing feature of the maize embryonic root system. These are roots formed endogenously from the scutellar node (region of the embryo between the primary root and young shoot) during embryogenesis after the primary root emerges (Feldman 1994). The number of seminal roots can vary widely among different genetic backgrounds, ranging from 0 to 13 per seedling (Feldman 1994; Hochholdinger 2009; Kiesselbach 1949). During the seed germination phase, most of the energy requirement is met by the kernel's nutrient reserves. However, the embryonic root system is crucial in water uptake and any damage to the root system can adversely affect seedling growth and development (Nielsen 2012). In maize, embryonic roots generally form the majority of the rootstock up to 2 weeks post-germination. Subsequently, nodal roots take over as the major constituent of the root system and, together with lateral roots, form the bulk of the adult plant root system (Hoppe et al. 1986).

Nodal roots and lateral roots constitute the postembryonic root system in maize. There are two classes of nodal roots in maize, the crown roots formed below ground and the brace roots formed aboveground. Both classes of shoot-borne roots are initiated in whorls around the stem (Hoppe et al. 1986). An average plant produces six whorls of crown roots and two to three whorls of brace roots, giving rise to approximately 70 nodal roots in total during its life cycle (Hoppe et al. 1986). These roots develop from primordia that originate from ground tissue cells (root cortical cells) opposite to collateral vascular bundles (Martin and Harris 1976). In addition to water and nutrient uptake, they also provide lodging resistance. Lateral roots are those roots formed from other root types (originate from primordia formed from pericycle and endodermis cells) and appear as branches of the main root (thus, also referred as branch roots) from which they originate. These roots are distinguishable from other root types by their relatively short lengths and often determinate apical meristems (Varney and McCully 1991). There can be several orders of lateral roots present in a root system, with the first series of lateral roots termed secondary roots, those borne from them termed tertiary roots, and so forth (Lynch 1995; Esau 1977). Lateral roots greatly increase the absorptive surface area of the root system and act as the main route for water and nutrient uptake (McCully and Canny 1988; Varney and Canny 1993; Postma et al. 2014b).

Different maize root types show comparable radial and longitudinal organization (reviewed by (Hochholdinger 2009)). Radially, root cells are organized in a concentric ring of layers starting from the outermost epidermis to the inner vascular tissue or stele (Hochholdinger et al. 2004b). A single layer of cells constitutes the epidermis. There are two kinds of cells in the epidermis, the root hair-forming trichoblasts and the atrichoblasts that do not form root hairs (Row and Reeder 1957). The ground tissue follows the epidermis and usually contains eight to 15 layers of cortex and a single layer of endodermis characterized by Casparian strips, which are localized lignified thickenings of the secondary cell wall along the medial plane of cells in this tissue (Hochholdinger 2009). The endodermis separates the inner vasculature from the outer root tissues and functions as a barrier to the radial flow of nutrients and water between the vascular cylinder and the outside environment (Dinneny 2014; Robbins et al. 2014). In the mature parts of the root, the outermost cortical cell layer known as the exodermis transforms into an additional barrier layer with a lignified and suberized cell wall and Casparian strips (Feldman 1994; Hose 2001). The innermost vascular tissue is composed of xylem and phloem cells and encircled by a layer of pericycle cells. A mature maize root shows a polyarch organization with a central protostele and many xylem arms (Kiesselbach 1949). The number of xylem vessels varies between root types, for example, when the primary root shows six to ten metaxylem vessels, a nodal root from a higher node may contain up to 48 metaxylem vessels (reviewed by (Hochholdinger 2009)). Phloem strands, which conduct photosynthates, are present in between the xylem vessels (Feldman 1994).

Longitudinally, the maize root can be divided into different zones (Hochholdinger 2009). Starting at the distal end is the root cap composed of up to

10,000 cells (Feldman 1994; Ishikawa and Evans 1993). The root cap cells are involved in a multitude of processes including sensing gravity and moisture (Feldman 1994; Ishikawa and Evans 1993). They also facilitate penetration of the root tip through the soil by secreting mucilage and acting as a cover to protect the meristematic region while the root grows. Proximal to the root cap is the QC composed of approximately 800–1200 cells (Jiang et al. 2003). Proximal to the QC are the rapidly dividing proximal meristem cells and their initials (Hochholdinger et al. 2004b). After several rounds of rapid cell division in the meristematic region, cells enter the elongation zone and undergo anisotropic expansion before they move into the next differentiation/maturation zone, characterized by the presence of root hairs, where cells of all tissues attain final shape and function (Ishikawa and Evans 1993).

### *11.2.2 Root System Development in Switchgrass*

Switchgrass, a perennial grass from the prairies of North America, has recently emerged as a frontrunner for the development of lignocellulosic biofuel crops (Lee 2006; Monti 2012). The embryonic root system consists of a primary root (Newman and Moser 1988). As the seedling develops, additional postembryonic roots emerge. Crown roots form on the lower stem nodes and several orders of lateral roots together form the bulk of the adult rootstock (Metcalfe and Nelson 1985). Switchgrass root systems show an exponential growth spurt in the initial 3 weeks of seedling growth and later gradually slow down (Dalrymple and Dwyer 1967). This active growth phase is thought to be critical for the successful growth and establishment of switchgrass (Zegada-Lizarazu et al. 2010). Another feature of the switchgrass root system is its ability to store nutrient resources in the long-lived root system and rhizomes (underground stems). Switchgrass can mobilize nutrient resources such as carbohydrates and nitrogen from the shoot system to the root system and vice versa depending on the growth season (Zegada-Lizarazu et al. 2010; Vogel 2004; Lemus et al. 2008, 2009; Wayman et al. 2014). This feature could be one of the reasons why switchgrass plants can grow on marginal lands but still maintain productivity (Parrish and Fike 2005). It has been noted that in mature switchgrass plants the upper 1-m of soil contains the bulk of the root system although the roots have been documented to reach a depth of 3 m or more (Ma et al. 2000).

### *11.2.3 Root System Development in Sugarcane*

Contrary to most other members of the Panicoideae grass family, sugarcane is commonly propagated asexually using pieces of the stem referred to as sett. Hence, the whole root system is postembryonic in origin consisting of shoot-borne sett roots,

shoot roots, and associated lateral roots (Smith et al. 2005). The first batch of roots formed is the sett roots, usually appearing within 24 h after planting the sett (Venkatraman and Thomas 1922; van Dillewijn 1922; Glover 1967). These fine roots are characterized by their high degree of branching and substitute for embryonic roots in the establishment of the new plant. However, as the plant becomes mature, sett roots gradually disappear (Glover 1967). After about a week post-planting, a new root type starts emerging from the sett base referred to as shoot roots (van Dillewijn 1922; Glover 1967). Shoot roots are thicker and grow more rapidly than sett roots and are the major root type in the sugarcane root system (Smith et al. 2005; van Dillewijn 1922; Glover 1967). Broadly, shoot roots can be categorized into three functional types: (1) Buttress roots are the first group of shoot roots produced, show relatively little branching and are critical in anchorage (initially growing outwards and then downwards into the soil) of the plant (Evans 1935), (2) Superficial roots are finer, highly branched roots produced from higher nodes above the stem base and are key players in water and nutrient uptake, (3) Rope roots are agglomerations of vertical roots that grow vertically into deep soil (often at depths of more than 6 m) and aid in extracting water from deep underwater reserves (Smith et al. 2005; Evans 1936). It has been reported that in modern sugarcane cultivars, the overall root system structure appears to be slightly different, for example, the rope roots are less prominent (Moore 1987). Sugarcane root systems are distinct in their high degree of plasticity, root length density, and distribution pattern in soil (more deeply distributed) from other crops (Smith et al. 2005; Blackburn 1984; Jackson et al. 1996). As with other grasses, the sugarcane root-to-shoot ratio is at its highest during early stages of plant growth and later on gradually declines as the plant matures (Smith et al. 2005).

### *11.2.4 Root System Development in Sorghum*

In Sorghum, the embryonic root system consists of a single primary root produced from the base of the embryo (Singh et al. 2010). The primary root functions throughout the plant life cycle, and, although similar to other grass species, its role is highly diminished after the onset of the postembryonic root system (Ernst 1948). About 1 week after germination, lateral roots start forming and subsequently, at the 4–5 leaf stage, the crown roots develop, which, together, form the postembryonic root system. As the plant grows, crown/nodal roots develop at regular intervals from sequential nodes beginning at the stem base throughout the life cycle (Singh et al. 2010). Nodal roots emerging from higher nodes (above ground) appear thicker and show a more vertical angle with respect to gravity than their counterparts from lower nodes. These roots are commonly referred to as buttress roots and once they enter the soil show a reduction in root diameter (Ernst 1948). It has been reported that, on average, Sorghum plants produce twice as many lateral roots as maize at any given stage of development and is thought to be a contributing factor for this species' high drought tolerance (Miller 1916).

## 11.3 Setaria as a Model for Panicoid Grass Root Systems

### *11.3.1 Setaria viridis: A Grass Genetic Model System*

Though there are reports describing the morphological characteristics and gene mutants affecting root type development in a few of the panicoid grasses, detailed analyses to dissect the underlying genetic regulatory networks that govern various aspects of root system dynamics/development such as cell-fate decisions and patterning steps are still missing (Bellini et al. 2014; Hochholdinger and Zimmermann 2008; Marcon et al. 2013). This is primarily due to the difficulties in systematically growing and studying these plants due to constraints such as large size, complex genetic makeup, lack of molecular biology tools, long life cycle, and general difficulty of studying root traits in soil. Availability of large numbers of gene mutants and other genetic resources (reporter lines, techniques to alter gene functions, etc.) are essential to characterizing the underlying molecular mechanisms regulating root growth and development in any plant system. In this context, a panicoid grass model species with all the ideal characteristics that one would expect for a genetic model system is of particular significance.

Recently, *Setaria viridis* has emerged as a potential genetic model system to study the panicoid subfamily of grasses. It has a relatively small sequenced genome, short life cycle, robust seed production, ease of growth under controlled conditions, and is transformable (Doust et al. 2009; Brutnell et al. 2010; Van Eck and Swartwood 2015). A method to break the long dormancy of freshly harvested seeds (often up to 4 months) has also been described by our group, thus enabling the full utilization of the rapid life cycle trait (Sebastian et al. 2014). In addition, genetic tools and resources are currently being developed in various labs across the world. As an added advantage, *S. viridis* is closely related to foxtail millet (*S. italica*) and the two species are intercrossable making it feasible to explore crop domestication with this system. Thus, Setaria offers tremendous potential in improving our understanding of root biology in this economically important group of plants.

### *11.3.2 Root System Development in Setaria*

*Setaria viridis* seeds germinate quickly and the primary root often becomes visible 24 h after imbibition on gel-based media (half-strength MS media, 0.5 % sucrose with 0.6 % Gelrite at 29 °C). The embryonic root system comprises a single primary root emerging from the basal pole of the embryo (Fig. 11.1d). Although occasionally some seedlings produce 1–2 seminal roots, they are mostly absent in Setaria. Around 3–4 days after planting (DAP), lateral roots start to emerge from the primary root (Fig. 11.1d). From seed germination to the emergence of crown roots, the sole primary root and associated lateral roots sustain seedling growth. The first crown roots emerge from the coleoptilar node (first shoot node) around 6–9 DAP and together with lateral roots make up the postembryonic root system in

Setaria (Fig. 11.1e and f). Both primary and crown roots can form several orders of lateral roots as the plant grows (Fig. 11.1c and f). Nadeau and Morrison reported that Setaria roots penetrate soil to a depth of nearly 60 cm with the highest concentration of roots at a depth of around 20–30 cm based on their field experiments (Nadeau and Morrison 1986). As with many other grasses, root system development in Setaria appears to be correlated with the production of leaves and tillers (Watt et al. 2009). For example, the first crown root usually appears when the plant is at the 3-leaf stage.

The primary root is separated from the crown and aerial tissues by the mesocotyl, an internode connecting the scutellar and coleoptilar nodes (Hoshikawa 1969). We have observed that the mesocotyl is more prominent in seeds that are germinated in soil compare to those seedlings that are germinated on tissue culture plates. This appears to be an adaptation to place the coleoptile at or near the soil surface irrespective of the seed planting depth. The mesocotyl can often produce adventitious roots, especially in soil-grown plants. It is remarkable to note that when the shoot system is just a few centimeters in length, the root system can be over three times as large in size; thus demonstrating the incredible rates of root growth these plants are capable of (data not shown). In Setaria, crown roots and lateral roots that are produced from a single plant at different time points often show variations in thickness, a phenomenon known as heteroblasty (Hou and Hill 2002; Zotz et al. 2011).

### *11.3.3 Cellular Organization of Setaria viridis Roots*

In seedlings that are germinated on tissue culture plates, the primary root is around 200–250 μm in diameter. It has a highly organized and radially symmetric cellular structure similar to other grass species (Fig. 11.2a–d). Longitudinally, the Setaria primary root can be subdivided into meristematic, elongation, and differentiation/maturation zones. The meristematic zone is characterized by rapidly proliferating meristematic cells. As the root grows, these cells are gradually displaced into the adjacent elongation zone, where they continue to elongate. The meristematic zone also harbors the stem cell niche and the QC. Further work is required to clearly identify the number of QC cells and the organization of the stem cell niche in Setaria roots. The apex of the meristematic zone is covered by the root cap containing >ten layers of collumella/calyptrogen cells (data not shown). Proximal to the elongation zone is the zone of differentiation/maturation, where the cells finally attain their predetermined shape, size, and cell fate and constitute the different tissue types of the root. Epidermal cells in the differentiation zone form root hairs; unicellular extensions of the epidermis which functions in the absorption of nutrients and water by increasing the overall root surface area (reviewed by (Mendrinna and Persson 2015)).

In the Setaria primary root, cells are arranged in a concentric ring of layers starting from the outermost epidermis (Fig. 11.2b and d). Three layers of cortical tissue develop interior to the epidermis with the outermost layer giving rise to the presumptive exodermis. A single layer of endodermis separates the cortex from

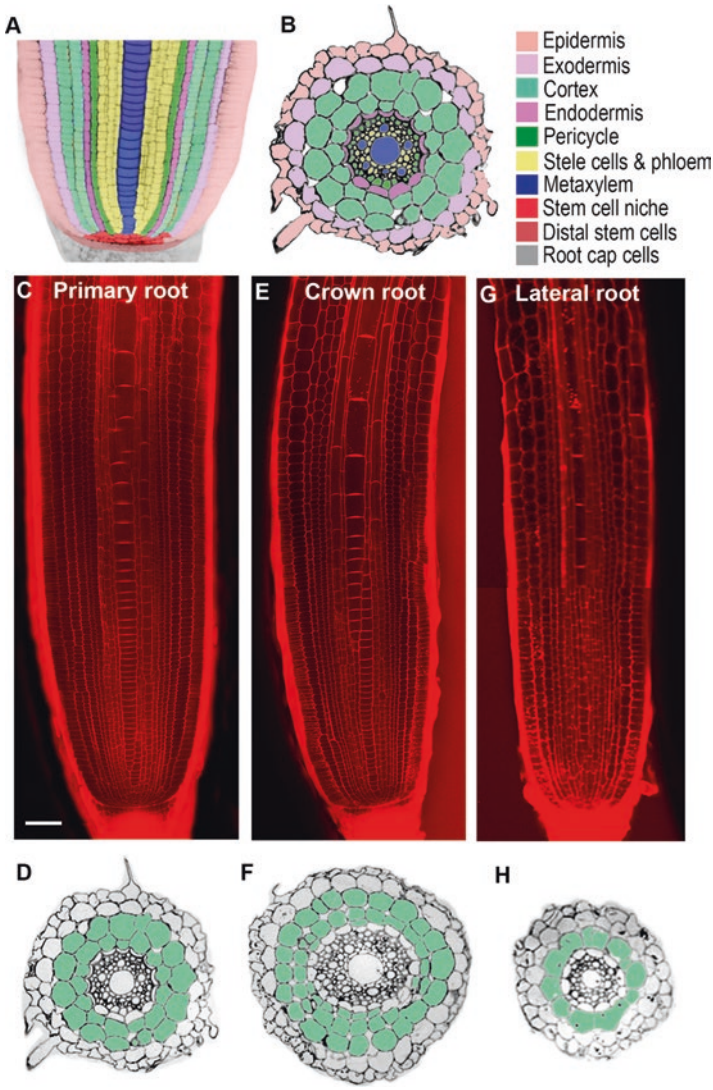


**Fig. 11.2** Cellular organization of *Setaria viridis* roots. (**a**, **b**) False-colored longitudinal (**a**) and radial (**b**) cross-section images of a Setaria root tip showing cell/tissue organization. Various colors indicate different cell/tissue types. (**c**, **e**, **g**) Longitudinal optical cross-sections of mPS-PI stained Setaria primary (**c**), crown (**e**), and lateral (**g**) roots. (**d**, **f**, **h**) Radial cross-section images of Setaria primary (**d**), crown (**f**), and lateral (**h**) roots obtained through thin sectioning. Scale bars: 55 μm

the vasculature. The innermost stele, which encompasses the xylem and phloem tissues, is encircled by a layer of pericycle cells (Fig. 11.2b and d). As in other panicoid grasses, the Setaria primary root stele has a polyarch organization, with varying numbers of central metaxylem (or late metaxylem) strands. Soon after emergence, roots have only a single central metaxylem strand, but as the root

matures, it may contain four or more strands. Surrounding these central xylem strands are the peripheral xylem strands. Depending on the age of the root, their number can also vary. At 4 DAP, the primary root has one central metaxylem strand and five to six peripheral xylem strands. Compared to the central xylem elements, the peripheral xylem elements are smaller in diameter (Fig. 11.2b and d). Phloem elements are found in between these peripheral xylem strands. As a characteristic feature of grass roots, the parenchymatous cells that are present between the xylem and phloem strands remain as such rather than differentiating into vascular cambium.

In Setaria, the crown roots are usually larger in diameter than the primary roots (250–300 μm) while lateral roots are the narrowest root type in the root system (100–150 μm in diameter). Anatomically, the crown roots and lateral roots are similar to the primary root (Fig. 11.2e–h); however, in certain instances, both crown and lateral roots show variation in the number of cortical cell layers (>two layers in crown roots and ≤2 in lateral roots).

## 11.4 Conclusions and Prospects

Although grasses have paramount importance to us both directly as a food source and indirectly as animal feed and a valuable source of sustainable clean energy, we are far behind in understanding their biology and growth mechanisms. This is especially true concerning grass root biology, which is particularly understudied and is paradoxically an area likely to hold potential for improvements in the overall productivity of our agriculture systems. This is largely due to the drawbacks grasses possess as model systems amenable to genetic studies and manipulations. However, with the emergence of new model species such as *Setaria viridis*, the tools available to understand grass root biology may enable a more mechanistic and comprehensive understanding of the process. Development of genetic resources such as high-quality fully annotated genomes, availability of genetic mutants, tissue/cell-type reporters, capabilities to edit the genome using the CRISPR-Cas9 system (Ran et al. 2013; Feng et al. 2013), and a rich and diverse germplasm collection with excellent SNP (single nucleotide polymorphism) data are all crucial in this pursuit. Creation of cell-type specific data sets that facilitate thorough characterization of molecular regulatory events at cellular resolution and novel methods to image the root system such as the recently developed GLO-Root imaging system (Rellán-Álvarez et al. 2015), which allows visualization of the root system under physiologically relevant conditions, are also important in dissecting the intricacies of grass root biology and environmental responses.

**Acknowledgments** We would like to thank Muh-ching Yee, Neil E Robbins, and Wei Feng for comments on the manuscript. Mandy Ka Wong for assistance in figure preparation. USDA (North Central Regional Plant Introduction Station, Iowa State University) and Tom Brutnell for *S. viridis* seeds. Funding is provided by a grant from the Department of Energy Biological and Environmental Research program (#DE-SC0008769) to J. R. D.

## References

Aiken RM, Smucker AJ. Root system regulation of whole plant growth. Annu Rev Phytopathol. 1996;34:325–46.

Bellini C, Pacurar DI, Perrone I. Adventitious roots and lateral roots: similarities and differences. Annu Rev Plant Biol. 2014;65:639–66.

Bennett T, Scheres B. In: Plant development. Amsterdam: Elsevier; 2010. p. 67–102.

Blackburn F. Sugar-cane. New York: Longman; 1984. 414 p. ISBN: 0- 582-46028-X.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, et al. Setaria viridis: a model for C4 photosynthesis. Plant Cell. 2010;22:2537–44.

Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P. Structure and functions of the bacterial microbiota of plants. Annu Rev Plant Biol. 2013;64:807–38.

Coudert Y, Périn C, Courtois B, Khong NG, Gantet P. Genetic control of root development in rice, the model cereal. Trends Plant Sci. 2010;15:219–26.

Dai X, Wang Y, Yang A, Zhang W-H. OsMYB2P-1, an R2R3 MYB transcription factor, is involved in the regulation of phosphate-starvation responses and root architecture in rice. Plant Physiol. 2012;159:169–83.

Dalrymple RL, Dwyer DD. Root and shoot growth of five range grasses. J Range Manage. 1967;20:141–5.

De Smet I, Signora L, Beeckman T, Inzé D, Foyer CH, et al. An abscisic acid-sensitive checkpoint in lateral root development of Arabidopsis. Plant J. 2003;33:543–55.

Dinneny JR. A gateway with a guard: how the endodermis regulates growth through hormone signaling. Plant Sci. 2014;214:14–9.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL. Foxtail millet: a sequence-driven grass model system. Plant Physiol. 2009;149:137–41.

Duan L, Dietrich D, Ng CH, Chan PMY, Bhalerao R, et al. Endodermal ABA signaling promotes lateral root quiescence during salt stress in Arabidopsis seedlings. Plant Cell. 2013;25:324–41.

Ernst A. Anatomy and morphology of the vegetative organs of Sorghum Vulgäre. U.S. DEPT Agric Tech Bull No 957. 1948.

Esau K. Anatomy of seed plants. 2nd ed. New York: Wiley; 1977.

Evans H. The root-system of the sugar-cane: I. Methods of study. Emp J Exp Agric. 1935;3:351–63.

Evans H. The root-system of the sugar-cane: II. Some typical root-systems. Emp J Exp Agric. 1936;4:208–21.

Feldman L. The maize root. In: Freeling M, Walbot V, editors. The maize handbook. New York, NY: Springer; 1994.

Feng Z, Zhang B, Ding W, Liu X, Yang D-L, et al. Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. 2013;23:1229–32.

Giussani LM, Cota-Sánchez JH, Zuloaga FO, Kellogg EA. A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C4 photosynthesis on JSTOR. Am J Bot. 2001;88:1993–2012.

Glover J. The simultaneous growth of sugarcane roots and tops in relation to soil and climate. Proc S Afr Sugar Technol Assoc. 1967;41:143–59.

GPWG. Phylogeny and subfamily classification of the grasses. Ann Mo Bot Gard. 2001;88:373–457.

Grossman JD, Rice KJ. Evolution of root plasticity responses to variation in soil nutrient distribution and concentration. Evol Appl. 2012;5:850–7.

Hammer GL, Dong Z, McLean G, Doherty A, Messina C, et al. Can changes in canopy and/or root system architecture explain historical maize yield trends in the U.S. corn belt? Crop Sci. 2009;49:299.

Hochholdinger F. The maize root system: morphology, anatomy, and genetics. In: Bennetzen JL, Hake SC, editors. Handbook of maize: its biology. New York, NY: Springer; 2009.

Hochholdinger F, Zimmermann R. Conserved and diverse mechanisms in root development. Curr Opin Plant Biol. 2008;11:70–4.
Hochholdinger F, Zimmermann R. Molecular and genetic dissection of cereal root system development. In: Beckman T, editor. Root development. Oxford: Wiley-Blackwell; 2009. p. 175–91.
Hochholdinger F, Park WJ, Sauer M, Woll K. From weeds to crops: genetic analysis of root development in cereals. Trends Plant Sci. 2004a;9:42–8.
Hochholdinger F, Woll K, Sauer M, Dembinsky D. Genetic dissection of root formation in maize (*Zea mays*) reveals root-type specific developmental programmes. Ann Bot. 2004b;93:359–68.
Hoppe DC, McCully ME, Wenzel CL. The nodal roots of Zea : their development in relation to structural features of the stem. Can J Bot. 1986;64:2524–37.
Hose E. The exodermis: a variable apoplastic barrier. J Exp Bot. 2001;52:2245–64.
Hoshikawa K. Underground organs of the seedlings and the systematics of Gramineae. Bot Gaz. 1969;130:192–203.
Hou G, Hill JP. Heteroblastic root development in Ceratopteris richardii (Parkeriaceae). Int J Plant Sci. 2002;163:341–51.
Hu B, Zhu C, Li F, Tang J, Wang Y, et al. LEAF TIP NECROSIS1 plays a pivotal role in the regulation of multiple phosphate starvation responses in rice. Plant Physiol. 2011;156:1101–15.
Ishikawa H, Evans ML. The role of the distal elongation zone in the response of maize roots to auxin and gravity. Plant Physiol. 1993;102:1203–10.
Itoh J-I, Nonomura K-I, Ikeda K, Yamaki S, Inukai Y, et al. Rice plant development: from zygote to spikelet. Plant Cell Physiol. 2005;46:23–47.
Jackson RB, Canadell J, Ehleringer JR, Mooney HA, Sala OE, et al. A global analysis of root distributions for terrestrial biomes. Oecologia. 1996;108:389–411.
Jiang K, Meng YL, Feldman LJ. Quiescent center formation in maize roots is associated with an auxin-regulated oxidizing environment. Development. 2003;130:1429–38.
Khanna M, Onal H, Chen X, Huang H. Meeting biofuels targets: implications for land use, greenhouse gas emissions and nitrogen use in Illinois. Environ Rural Dev Impacts Conf Oct 15–16, 2008, St Louis, Missouri. 2008. http://ideas.repec.org/p/ags/fftren/53491.html
Kiesselbach TA. The root system. In: The structure and reproduction of corn. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1949.
Kong X, Zhang M, De Smet I, Ding Z. Designer crops: optimal root system architecture for nutrient acquisition. Trends Biotechnol. 2014;32:597–8.
Lee R. Switchgrass as a bioenergy crop. A Publ ATTRA—Natl Sustain Agric Inf Serv, 1-800-346-9140. 2006. www.attra.ncat.org
Lemus R, Parrish DJ, Abaye O. Nitrogen-use dynamics in switchgrass grown for biomass. Bioenerg Res. 2008;1:153–62.
Lemus R, Parrish DJ, Wolf DD. Nutrient uptake by “Alamo” switchgrass used as an energy crop. Bioenerg Res. 2009;2:37–50.
Lewandowski I, Scurlock JMO, Lindvall E, Christou M. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. Biomass Bioenerg. 2003;25:335–61.
Lynch J. Root architecture and plant productivity. Plant Physiol. 1995;109:7–13.
Ma Z, Wood CW, Bransby DI. Impacts of soil management on root characteristics of switchgrass. Biomass Bioenerg. 2000;18:105–12.
Marcon C, Paschold A, Hochholdinger F. Genetic control of root organogenesis in cereals. Methods Mol Biol. 2013;959:69–81.
Martin EM, Harris W. Adventitious root development from the coleoptilar node in Zea mays. L. Am J Bot. 1976;63:890–7.
McCully ME, Canny MJ. Pathways and processes of water and nutrient movement in roots. Plant Soil. 1988;111:159–70.
Mendrinna A, Persson S. Root hair growth: it’s a one way street. F1000Prime Rep. 2015;7:23.
Metcalfe DS, Nelson CJ. The botany of grasses and legumes. In: Heath ME et al., editors. Forages. 4th ed. Ames: Iowa State Univ Press; 1985. p. 52–63.

Miller EC. Comparative study of the root systems and leaf areas of corn and the Sorghums. J Agric Res. 1916;6:311–32.
Monti A, editor. Switchgrass—a valuable biomass crop for energy. Berlin: Springer; 2012.
Moore PH. Anatomy and morphology. In: Heinz DJ, editor. Sugarcane improvement through breeding. Amsterdam: Elsevier; 1987. p. 85–142.
Nadeau LB, Morrison IN. Influence of soil moisture on shoot and root growth of green and yellow foxtail (Setaria viridis and S. lutescens. Weed Sci. 1986;34:225–32.
Newman PR, Moser LE. Seedling root development and morphology of cool-season and warm-season forage grasses. Agron Hortic—Fac Publ. 1988. http://digitalcommons.unl.edu/agronomyfacpub/82
Nielsen R. Early-planted corn and cold weather. Corny News Network, Purdue Ext. 2012. http://www.kingcorn.org/news/articles.12/EarlyCornColdWthr-0412.html
Onderdonk JJ, Ketcheson JW. A standardization of terminology for the morphologocal description of corn seedlings. Can J Plant Sci. 1972;52:1003–6.
Orman-Ligeza B, Parizot B, Gantet PP, Beeckman T, Bennett MJ, et al. Post-embryonic root organogenesis in cereals: branching out from model plants. Trends Plant Sci. 2013;18:459–67.
Osborne CP, Salomaa A, Kluyver TA, Visser V, Kellogg EA, et al. A global database of C4 photosynthesis in grasses. New Phytol. 2014;204:441–6.
Parrish DJ, Fike JH. The biology and agronomy of switchgrass for biofuels. CRC Crit Rev Plant Sci. 2005;24:423–59.
Petricka JJ, Winter CM, Benfey PN. Control of Arabidopsis root development. Annu Rev Plant Biol. 2012;63:563–90.
Postma JA, Schurr U, Fiorani F. Dynamic root growth and architecture responses to limiting nutrient availability: linking physiological models and experimentation. Biotechnol Adv. 2014a;32:53–65.
Postma JA, Dathe A, Lynch JP. The optimal lateral root branching density for maize depends on nitrogen and phosphorus availability. Plant Physiol. 2014b;166:590–602.
Pray CE, Nagarajan L. Innovation and Research by Private Agribusiness in India. IFPRI Discuss Pap 01181, 2002.
Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, et al. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281–308.
Rellán-Álvarez R, Lobet G, Lindner H, Pradier P-LM, Sebastian J, et al. Multidimensional mapping of root responses to soil environmental cues using a luminescence-based imaging system. eLife. 2015;19:4. doi:10.7554/eLife.07597.
Robbins NE, Trontin C, Duan L, Dinneny JR. Beyond the barrier: communication in the root through the endodermis. Plant Physiol. 2014;166:551–9.
Rogers ED, Benfey PN. Regulation of plant root system architecture: implications for crop advancement. Curr Opin Biotechnol. 2015;32:93–8.
Row HC, Reeder JR. Root-hair development as evidence of relationships among genera of gramineae. Am J Bot. 1957;44:596–601.
Scarpella E, Meijer AH. Pattern formation in the vascular system of monocot and dicot plant species. New Phytol. 2004;164:209–42.
Sebastian J, Lee JY. In: Root apical meristems. Chichester: Wiley; 2013. www.els.net. doi:10.1002/9780470015902.a0020121.pub2
Sebastian J, Wong MK, Tang E, Dinneny JR. Methods to promote germination of dormant Setaria viridis seeds. PLoS One. 2014;9:e95109.
Singh V, van Oosterom EJ, Jordan DR, Messina CD, Cooper M, et al. Morphological and architectural development of root systems in sorghum and maize. Plant Soil. 2010;333:287–99.
Smith DM, Inman-Bamber NG, Thorburn PJ. Growth and function of the sugarcane root system. Field Crop Res. 2005;92:169–83.
van Dillewijn C. Botany of sugarcane. Waltham, MA: Chronica Botanics; 1922. 371 p.
Van Eck J, Swartwood K. Setaria viridis. Methods Mol Biol. 2015;1223:57–67.
Varney GT, Canny MJ. Rates of water uptake into the mature root system of maize plants. New Phytol. 1993;123:775–86.

Varney GT, McCully ME. The branch roots of Zea. II. Developmental loss of the apical meristem in field-grown roots. New Phytol. 1991;118:535–46.
Venkatraman TS, Thomas R. Sugarcane root systems: studies in development and anatomy. Agric J India. 1922;17:381–8.
Vogel KP. Switchgrass. In: Sollenberger LE, Moser LE, Burson BL, editors. Warm-season (c4) grasses. Agron monogr 45. Madison, WI: ASA, CSSA, SSSA; 2004.
Watt M, Schneebeli K, Dong P, Wilson IW. The shoot and root growth of Brachypodium and its potential as a model for wheat and other cereal crops. Funct Plant Biol. 2009;36:960.
Wayman S, Bowden RD, Mitchell RB. Seasonal changes in shoot and root nitrogen distribution in switchgrass (Panicum virgatum). Bioenerg Res. 2014;7:243–52.
White PJ, George TS, Gregory PJ, Bengough AG, Hallett PD, et al. Matching roots to their environment. Ann Bot. 2013;112:207–22.
Zegada-Lizarazu W, Elbersen HW, Cosentino SL, Zatta A, Alexopoulou E, et al. Agronomic aspects of future energy crops in Europe. Biofuels Bioprod Bior. 2010;4:674–91.
Zhou J, Jiao F, Wu Z, Li Y, Wang X, et al. OsPHR2 is involved in phosphate-starvation signaling and excessive phosphate accumulation in shoots of plants. Plant Physiol. 2008;146:1673–86.
Zhu X-G, Long SP, Ort DR. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Curr Opin Biotechnol. 2008;19:153–9.
Zotz G, Wilhelm K, Annette B. Heteroblasty—a review. Bot Rev. 2011;77:109–51. doi:10.1007/s12229-010-9062-8.

# Part IV
# Genetics

# Chapter 12
# The Effect of Photoperiod on Flowering Time, Plant Architecture, and Biomass in Setaria

**Andrew N. Doust**

**Abstract** The effect of photoperiods of 8 h (8:16 light:dark), 12 h (12:12), and 16 h (16:8) on flowering time, plant architecture, and biomass production were investigated in an RIL population derived from a cross between domesticated foxtail millet (*Setaria italica*) and its wild progenitor green foxtail (*S. viridis*). Flowering time, height, and biomass were found to be highly and positively correlated in all three photoperiod regimes. Branching, however, is weakly and variably associated with the other three traits. After the effects of variation in daily radiation and temperature were removed, ANOVA analyses of *Photoperiod* and *RIL* (genotype) found both factors and their interaction significant for all traits, with *RIL* and *Photoperiod * RIL* also explaining large amounts of variation. However, while *Photoperiod* by itself explained much of the variation in flowering time and in branching, it explained little of that for height and biomass. Regions were identified where all three trials identify QTL in the same genomic regions as well as QTL found in either the 8 and 12 h trials or the 12 and 16 h trials. This pattern may be evidence for differences in regulation between shorter and longer photoperiods. Comparison of QTL with previous greenhouse and field trials finds several overlapping QTL and multiple independent QTL. A well-supported QTL region on chromosome IV has been shown previously to contain a number of genes in the CONSTANS—FT pathway, and these results suggest that this pathway is conserved across photoperiods. Further genetic analysis of the multiple non-overlapping QTL regions between the photoperiod trials will be necessary to narrow down a list of candidate genes responsible for differences in flowering time and architecture between photoperiods.

**Keywords** Flowering time • Setaria • Photoperiod • Branching • Height • Biomass • QTL analysis • Foxtail millet • Green foxtail

A.N. Doust (✉)
Department of Plant Biology, Ecology and Evolution,
Oklahoma State University, Stillwater, OK 74078, USA
e-mail: andrew.doust@okstate.edu



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_12

## 12.1 Introduction

The potential of Setaria as a model system is primarily based on its attributes for genetic analysis, particularly the small diploid genome, small physical stature, $C_4$ photosynthetic capability, transformability, and a growing list of genetic and genomic resources (Bennetzen et al. 2012; Doust et al. 2009; Li and Brutnell 2011). Although there are other model grasses, including rice (the first sequenced grass genome) and Brachypodium (a wild grass in the pooid clade related to wheat, barley, and rye), Setaria has the advantage of being a $C_4$ grass in the panicoid clade, close to maize, sorghum, switchgrass, and pearl millet. In addition, there is substantial genetic and phenotypic differentiation amongst wild populations of green foxtail (*S. viridis*) as well as genetic changes associated with domestication in its domesticated variant, foxtail millet (*S. italica*). The potential of Setaria to be a new model system is especially significant because there is a high efficiency callus transformation system (Van Eck and Swartwood 2015) (Chap. 20), as well as recent reports on the success of spike dip transformation with *Agrobacterium*, which is the first for any grass system (Saha and Blumwald 2016) (Chap. 21).

The use of Setaria as a model for biofuel grasses has prompted interest in the genetic regulation of, and correlation between, traits such as flowering time, plant architecture, and biomass (Mauro-Herrera and Doust 2016; Mauro-Herrera et al. 2013; Doust et al. 2004). The wide latitudinal spread of *S. viridis* from high to subtropical latitudes in both hemispheres suggests that changes in photoperiod may have significant effects on these traits. In addition, Setaria appears to differ from other model grass systems in that the center of diversity of green foxtail and the domestication of foxtail millet from green foxtail appears to have occurred at a relatively high latitude (Jia et al. 2013) (Chaps 2, 3, and 4), raising the possibility that photoperiodic control of flowering in Setaria may not conform to the model that has emerged from rice, a species that evolved and was domesticated in the tropics (Vaughan et al. 2008). Photoperiodic control of flowering in pooid crops, such as wheat, barley, and the model species Brachypodium, differs from that in Setaria, because they require a vernalization response to achieve competency to flower (Higgins et al. 2010), a strategy not known in Setaria or other panicoid grasses.

In this chapter, the response of Setaria to changes in photoperiod is explored, using a recombinant inbred line (RIL) population derived from a cross between domesticated foxtail millet (*Setaria italica*) and its wild progenitor green foxtail (*S. viridis*) (Bennetzen et al. 2012). Much of what is known about the response of grasses to differences in photoperiod is from studies in rice, where there are two photoperiod-dependent pathways; one of these is homologous to that found in Arabidopsis and other land plants, while the other appears to be confined to grasses (Mauro-Herrera et al. 2013). The first pathway involves the key regulator CONSTANS, which positively regulates FT in Arabidopsis but whose ortholog, *HD1*, in rice negatively regulates the FT co-orthologs, *HD3A* and *RFT1*, active under short-day and long-day conditions, respectively (Hayama et al. 2003; Izawa et al. 2002; Song et al. 2010; Komiya et al. 2009). It is not known whether Setaria exhibits a long and short day signaling pathway in the same way that rice does,

although there are three co-orthologs of FT in the Setaria genome (Bennetzen et al. 2012). In contrast to rice, the functional FT homolog in maize (*ZCN8*) and its equivalent in sorghum are in a different clade of PEBP proteins (Lazakis et al. 2011; Meng et al. 2011; Wolabu et al. 2016). RNA-seq and qRT-PCR data from *S. viridis* (unpublished) suggests that one of the Setaria co-orthologs of FT as well as a Setaria homolog of *ZCN8* are expressed at the same time during the transition of the vegetative shoot apical meristem to an inflorescence meristem.

The second photoperiod-controlled flowering time pathway identified in rice involves the negative regulators *GHD7* and *EHD1*, which work together to precisely determine the length of photoperiod that will induce flowering. Homologs of these two genes have been identified as involved in flowering time regulation in maize and sorghum, suggesting that the *GHD7-EHD1* pathway is a grass-specific flowering time pathway (Hung et al. 2012; Murphy et al. 2011, 2014; Yang et al. 2014). Orthologs of *GHD7* and *EHD1* have been identified in the Setaria genome but not functionally tested.

The effect of photoperiod on traits that interact with flowering time, such as biomass accumulation and plant architecture, have been little studied in most panicoid grasses. However, information on the effect of photoperiod on plant growth is important for Setaria as a model system because it has not been selected for photoperiod insensitivity, unlike modern cultivars of maize and sorghum. A common growth chamber strategy is to grow it under a 12:12 h light:dark photoperiod regime, as will be seen in other investigations presented in this book (Chaps 10, 11, 13, 14, 18–21). Such a strategy minimizes the effect of environmental variation on phenotype and encourages rapid flowering and fast cycling of generations—important criteria for a model system. However, field grown *S. viridis* and *S. italica* are rarely grown under less than 14 h light, and may be grown in as much as 16 h light in higher latitudes. In photoperiod-sensitive plants, such as Setaria, these differences might be expected to produce differences in both flowering time and plant growth traits.

To investigate these questions, we have grown a RIL population derived from a cross between *Setaria italica* (foxtail millet) and *S. viridis* (green millet) in three different photoperiod regimes (8:16, 12:12, and 16:8 h light:dark), while minimizing variation in other environmental variables. We report here on a QTL analysis of variation in flowering time, plant architecture, and biomass under these photoperiod regimes and compare results with previously published analyses using the same RIL population in greenhouse and field environments (Mauro-Herrera and Doust 2016; Mauro-Herrera et al. 2013).

## 12.2 Materials and Methods

### *12.2.1 Plant Materials, Experimental Design, and Phenotyping*

A total of 182 $F_7$ RILs from an interspecific cross between *S. italica* accession B100×*S. viridis* accession A10 (Bennetzen et al. 2012) were evaluated for flowering time, plant height, total branching, and biomass at flowering in a walk-in growth

chamber at Oklahoma State University (Stillwater, OK). Three trials were undertaken, at photoperiod ratios (light:dark) of 8:16, 12:12, and 16:8. The chamber was kept at 30 % humidity and day and night temperatures were 28 and 22 °C, respectively. Two other variables, besides photoperiod duration, varied between trials. These were amount of daily radiation received (directly related to photoperiod duration) and temperature (as the combination of different day lengths and the difference in day and night temperatures led to differences in the average temperature of each trial). The effects of daily radiation and temperature cannot be separated in this study, and their values were 8.64 E (Einstein = mol $m^{-2}$ $s^{-1}$) and 24 °C in the 8 h trial, 12.96 E and 25 °C in the 12 h trial, and 17.28 E and 26 °C in the 16 h trial. Illumination from full spectrum fluorescent tubes averaged 300 μmol $m^{-2}$ $s^{-1}$. Three replicate pots of each RIL were grown in each experiment, with each pot having a single plant. Pots were randomized, and plants were spaced 8.5 cm apart. Pot volume was approximately 215 $cm^3$, and pots were filled with Metro-Mix 366 (Sun Gro Horticulture Canada Ltd). Plants were irrigated as needed with an aqueous complete fertilizer mix (Jack's mix: Nitrogen, Phosphorous and Potassium (20-20-20), JR Peters, PA).

### *12.2.2 Phenotypic Measurement*

We used days to heading as the measurement of flowering, with plants recorded as flowering when the inflorescence on the main culm was first visible in the sheath of the flag leaf (Mauro-Herrera et al. 2013). Culm height (height of the main stem of the grass plant) was measured from the base of the plant to the ligule (leaf collar) of the flag leaf on the main culm. Total branches comprised both tillers (at base of plant) and any aerial branches. Total aboveground biomass was measured by drying whole plants for at least 1–2 weeks in a plant drier, and then weighing after removing the roots.

### *12.2.3 Statistical Analyses*

Traits were tested for normality and transformed where appropriate. Relationships between traits were explored by bivariate Pearson phenotypic correlations, using both original variables (transformed where necessary) and with effect of RIL removed (by using the residuals obtained from an ANOVA for each trait with RIL as the independent variable). Boxplots of each parent and for the combined RIL population were made for each trait. Trait differences between photoperiods were analyzed using ANOVA. Because it is likely that days to flowering is affected by both photoperiod (measured by a plant as the length of darkness in each 24 h period) and carbon gain (directly related to hours of light and temperature), trait values were first regressed against the total amount of illumination each plant received until flowering, and residuals used in the ANOVA analyses. The model fitted for all ANOVA analyses consisted of two factors, Photoperiod (fixed) and RIL (random).

Partial eta squared values were calculated to estimate proportion of trait variance explained by each factor or interaction. All analyses were performed with SPSS version 21 (IBM SPSS, Armonk, NY).

### *12.2.4 QTL Analyses*

For QTL analyses, we used the previously published 684 marker genetic map (Mauro-Herrera et al. 2013). QTL Cartographer Unix version 1.16 (Basten et al. 1994, 2002) was used for QTL analyses with the composite interval mapping (CIM) method, a genome scan interval of 1 cm, a window size of 10, and the forward and backward regression method (Jansen and Stam 1994; Zeng 1994). QTL analyses were conducted for each trait in each photoperiod trial, as well as a joint analysis for each trait across all three trials. The joint analysis measured both main effect QTL detected across trials as well as QTL that had a significant genotype by trial interaction. LOD threshold values were estimated via 1000 permutations (Churchill and Doerge 1994; Doerge and Churchill 1996). Comparisons amongst the growth chamber trials and between growth chamber and previous greenhouse and field trials were conducted by comparing overlap between QTLs for each trait, especially with respect to the position of the maximum LOD values.

## 12.3 Results

### *12.3.1 Phenotypic Variation*

Trait distributions were tested for normality, and biomass and branch number were square root transformed to improve the normality of their distributions. Transformed trait values were used for these two traits in subsequent analyses.

In all three trials, *S. viridis* flowered before *S. italica*, with most of the RILs flowering at intermediate times (Fig. 12.1). Flowering time in the RILs was skewed towards that of the earlier flowering *S. viridis* plants. Flowering time of the *S. italica* plants was especially long and variable in the 16 h trial (Fig. 12.1). There was little transgressive segregation for flowering time. *S. viridis* plants were shorter than *S. italica* plants at flowering, and the RILs in general had plant heights skewed towards *S. viridis* (Fig. 12.1). However, there was substantial transgressive segregation for height in the RIL population at all photoperiods, with greatest transgressive segregation in the 16 h trial. Transgressive variation was also seen for total branch number in all trials, and the *S. viridis* parent always had more branches than the *S. italica* parent (Fig. 12.1). In the 8 and 12 h trials, the *S. italica* plants did not produce any tillers or aerial branches at all. *Setaria viridis* always had less biomass than *S. italica*, and the biomass of the RIL lines was skewed towards *S. viridis* (Fig. 12.1). There was some transgressive segregation for biomass, especially in the 16 h trial.

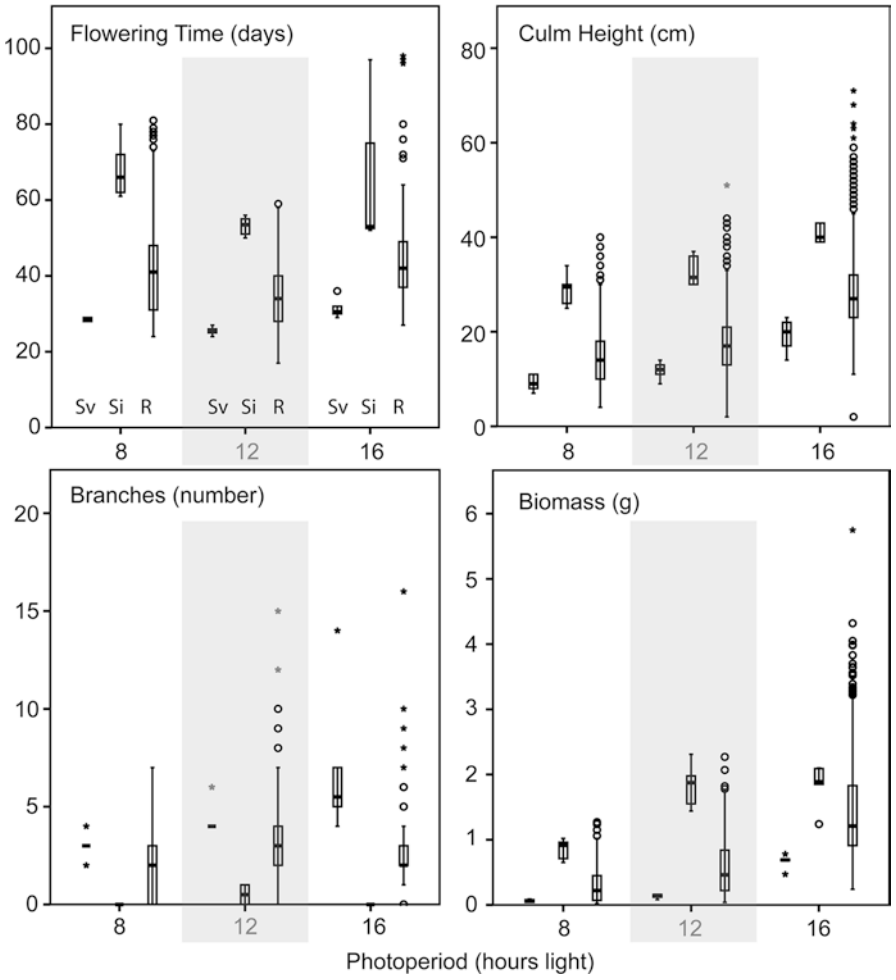


**Fig. 12.1** Boxplots of trait values for each trait in each of the three photoperiod regimes. All boxplots show the distribution for each parent and for the RIL population (Sv = *S. viridis*, Si = *S. italica*, R = RIL population)

Height, branching, and biomass showed a generally positive response to increasing length of photoperiod, for both parents and RILs. The same was not true for flowering time, where the 12 h trial exhibited the shortest flowering times, followed by the 8 h and then the 16 h.

The phenotypic traits in each individual trial showed high positive correlations between flowering time, height, and biomass (Table 12.1); the correlations were also found when the effect of RIL (genotype) was removed. The relationship between branching and the other three variables was less consistent although in all but one comparison the relationship between height and branching was significantly negative. The correlation between branching and flowering time varied from trial to

**Table 12.1** Correlations between traits in each of the photoperiod trials

| | | Trait values | | | Residuals[a] | | |
|---|---|---|---|---|---|---|---|
| Photoperiod | Trait | Culm height | Total branches[b] | Biomass[b] | Culm height | Total branches[b] | Biomass[b] |
| 8 h | Flowering time | ++ | ns | ++ | ++ | + | ++ |
| 8 h | Culm height | | −− | ++ | | ns | ++ |
| 8 h | Total branches[b] | | | ns | | | ++ |
| 12 h | Flowering time | ++ | −− | ++ | ++ | ns | ++ |
| 12 h | Culm height | | −− | ++ | | −− | ++ |
| 12 h | Total branches[b] | | | − | | | ns |
| 16 h | Flowering time | ++ | ns | ++ | ++ | ++ | ++ |
| 16 h | Culm height | | −− | ++ | | −− | ++ |
| 16 h | Total branches[b] | | | ns | | | ++ |

[a]Residuals have the effect of genotype (RIL) removed
[b]The values for these traits have been square root transformed
Positive significant correlations: + = $P<0.05$, ++ = $P<0.01$
Negative significant correlations: − = $P<0.05$, −− = $P<0.01$
Nonsignificant correlations = ns

trial, both with and without the effect of genotype (RIL) (Table 12.1). Correlations of total amount of light received with each of the traits were significant and positive and explained 56 % of the variation in flowering time, 48 % in height, 78 % in biomass, but only 2 % in branching.

The effect of photoperiod on each of the four traits was analyzed in the parents of the population by ANOVAs with and without the effect of total amount of illumination received. *Setaria viridis* was more sensitive to photoperiod changes than *S. italica* for both flowering time (*S. viridis* $p<0.001$, *S. italica* not significant) and branching (*S. viridis* $p<0.001$, *S. italica* $p<0.05$). Both *S. viridis* and *S. italica* had highly significant differences in height and biomass across photoperiods.

The effect of photoperiod and RIL genotype on each of the four traits was analyzed with ANOVAs using the residuals from a regression of the trait values against the total amount of illumination received. There were highly significant differences amongst both Photoperiod and RIL, and for the interaction between them. However, the amount of variation explained by each factor (partial eta squared values) and their interaction varied between traits (Table 12.2). RIL and Photoperiod * RIL explained large proportions of the variance for all traits, but Photoperiod by itself only explained large proportions of the variance for flowering time and for

**Table 12.2** Partial eta squared values for the ANOVA using the residuals of the four traits (after removing the effect of total amount of illumination received), showing the degree to which each factor explains variation in the traits

| Source factor | Flowering time | Height | Biomass (sqrt) | Branching (sqrt) |
|---|---|---|---|---|
| Photoperiod | 0.68 | 0.07 | 0.13 | 0.38 |
| RIL | 0.61 | 0.69 | 0.6 | 0.73 |
| Photoperiod * RIL | 0.85 | 0.56 | 0.53 | 0.43 |

All factors were significant for all traits
*Note*: Because RIL is a random sampling of all possible genotypes it is treated as a random factor. Therefore, the mean square used as an error term for the Photoperiod and RIL comparisons is the mean square for Photoperiod * RIL, and that for Photoperiod * RIL is the error mean square

branching, and very little of the variation in height and biomass. This suggests that the main driver for height and biomass is carbon gain driven by the number of illumination hours rather than photoperiod length.

### *12.3.2 QTL Analyses*

Across all individual trials ten QTL regions for flowering were identified, with five QTL in both the 8 and 12 h trials, and three in the 16 h trial. Three of the genomic regions contained QTL from multiple trials, these being on chromosomes IV and VII (8 and 12 h), and on chromosome VIII (12 and 16 h) (Fig. 12.2). There were 11 joint main effect QTL and six GxE effect QTL, indicating that the control of flowering has a significant environmental component. Eight QTL regions identified in individual trials overlapped with either a main or GxE QTL of the joint analysis. However, four of the joint QTL did not align with any of the individual QTL.

Across all individual trials nine QTL regions for height were identified, with three in the 8 h, five in the 12 h, and five in the 16 h trial, with only three genomic regions where QTL maximum LOD positions overlapped. These were on chromosomes IV (8 and 12 h), V (12 and 16 h), and IX (all three trials). Six main effect and four GxE effect QTL were identified in the joint analysis, of which one on chromosome IV overlaps with the 8 and 12 h trials, one on chromosome V overlaps with the 12 h trial, and one on chromosome IX that overlaps with all three trials.

Across all individual trials 11 QTL regions for biomass were identified, with five in the 8 h trial, five in the 12 h, and six in the 16 h trial. Four genomic regions contained overlapping QTL from the individual trials, these were on chromosomes IV (8, 12, and 16 h), V (8 and 12 h), VIII (12 and 16 h), and IX (8 and 12 h). There were eight joint main effect and six GxE effect QTL, of which four overlapped with QTL from the individual trials, on chromosomes II, III, IV, and V.

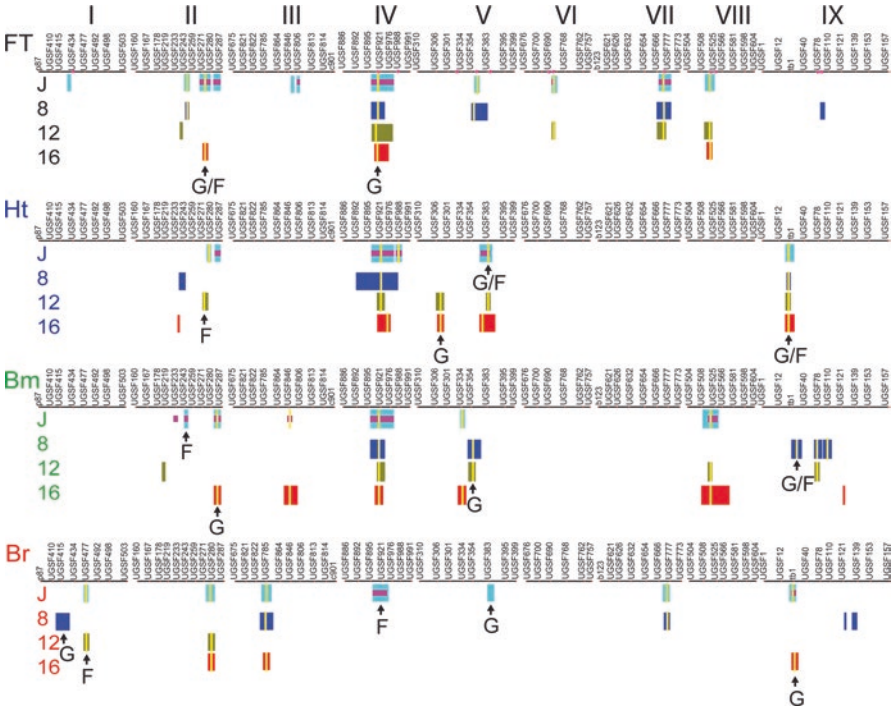


**Fig. 12.2** QTL map showing the distribution of QTL for each of the traits in each of the photoperiod environments (8 h—*dark blue*, 12 h—*ochre*, 16 h—*red*, as well as joint main QTL calculated for each trait across the three environments (*light blue*) and QTL by environment effects (*mauve* bars, often nested within *light blue* joint main QTL). G (greenhouse) and F (field) refer to regions where QTL from greenhouse and field trials (Mauro-Herrera and Doust 2016; Mauro-Herrera et al. 2013) overlap with QTL from this study

Across all individual trials eight QTL regions for branching were identified, with five QTL in the 8 h trial, two in the 12 h, and three in the 16 h trial. QTL overlapped in two genomic regions, the 12 and 16 h on chromosome II and the 8 and 16 h on chromosome III. There were seven joint main effect and two GxE effect QTL identified, of which five overlapped with individual trials. These were on chromosomes I (with 12 h), II (with 12 and 16 h), III (with 8 and 16 h), VII (with 8 h), and IX (with 16 h).

QTL for flowering time, height, and biomass show a striking overlap, especially on chromosome IV. Generally speaking, approximately one third of the QTL positions identified across the three trials were found in more than one trial (Table 12.3). In all but three of these regions, a joint QTL was also found, the exceptions being height on chromosome V (12 and 16 h), biomass on chromosome V (8 and 12 h), and biomass on chromosome IX (8 and 12 h). There was a greater percentage of overlap between individual trial QTL and joint QTL, as would be expected considering that the data from each individual trial contributes to the joint analysis (Table 12.3).

**Table 12.3** Average percentages of shared QTL between the different trials

| Trait | Amongst growth chamber individual trials[a] | QTL from individual trials that overlap with joint QTL[b] | QTL from greenhouse that overlap with growth chamber QTL[c] | QTL from field that overlap with growth chamber QTL[d] | Greenhouse versus field[e] |
|---|---|---|---|---|---|
| Flowering time | 3/10 | 7/12 | 2/8 | 1/5 | 4/9 |
| Height | 3/9 | 3/6 | 3/10 | 3/10 | 5/15 |
| Biomass | 4/11 | 3/7 | 3/12 | 2/10 | 5/17 |
| Branching | 2/8 | 5/7 | 3/10 | 2/5 | 1/14 |
| Mean ± S.D. | 0.31 ± 0.05 | 0.45 ± 0.12 | 0.28 ± 0.03 | 0.28 ± 0.10 | 0.28 ± 0.16 |

[a]Numerator is number of overlapping QTL between individual trials, denominator is total number of regions identified. Overlapping QTL can be in all three trials or in just two of the trials
[b]Numerator is number of overlapping QTL between individual trials and the joint analysis, denominator is total number of regions identified in the joint analysis
[c]Numerator is number of overlapping QTL between greenhouse QTL and individual + joint growth chamber analyses, denominator is total number of individual + joint growth chamber QTL identified
[d]Numerator is number of overlapping QTL between field QTL and individual + joint growth chamber analyses, denominator is total number of individual + joint growth chamber QTL identified
[e]Numerator is number of overlapping QTL between greenhouse and field, denominator is total number of regions identified in greenhouse and field

### *12.3.3 Comparisons of Growth Chamber Trials with Previous Greenhouse and the Field Trials*

QTL identified in the growth chamber trials were compared with those discovered in previous greenhouse and field trials (Mauro-Herrera et al. 2013, Mauro-Herrera and Doust 2016). The percentage of overlap of QTL between these different environments was similar to that between the individual growth chamber trials, and between greenhouse and field trials (Table 12.3). The region on chromosome IV that was significantly correlated with flowering time, height, and biomass in all three growth chamber photoperiod trials was also found for flowering time in the greenhouse and for branching in the field trial. Other QTL that were found in more than one growth chamber trial and in either or both of the greenhouse and field trials include those on chromosomes VII and VIII for flowering time. However, QTL from multiple trials for biomass on chromosome VIII and for branch number on chromosomes II and III in the growth chamber trials were not found in the greenhouse or field trials.

## 12.4 Discussion

Two main trends are seen in these trials. One is related to total amount of illumination received (although confounded with variation in average temperature) while the other is related to the duration of light and dark intervals. The correlations and

boxplots show that architectural and biomass traits have a positive response to increasing length of photoperiod, whereas the 12 h photoperiod regime gave the shortest flowering time, followed by the 8 h and then the 16 h regime. This may have been because the 12 h regime was the shortest viable photoperiod in terms of light quanta received for this C4 plant, and that 8 h light per day was simply not enough to allow flowering quickly. This is being tested in further experiments underway in our lab. The much longer time to flowering under the 16 h regime suggests that Setaria should be considered a facultative short day plant.

The positive response of height, branching and biomass to increasing photoperiod most probably reflects both the increase in light (and temperature) received each day as well as the increase in number of days to flowering that allows plants to continue to grow for a longer time period. This relationship is supported by the finding that flowering time is significantly correlated with height and biomass in all three trials, with or without the effect of genotype. When the effect of genotype is not considered, there is a significant positive correlation between flowering time and branch number in two of the three photoperiod regimes, but, when the effect of genotype is included, the relationship between branch number and flowering time is not significant, suggesting that different genotypes perform differently in different photoperiod regimes.

In the ANOVA analyses, we chose to concentrate on photoperiod and genotype (RIL), by eliminating confounding variation due to different levels of light intensity and/or temperature due to the different photoperiod lengths. ANOVA analyses of the four traits showed that Photoperiod, RIL, and Photoperiod * RIL explained significant proportions of the variance of all four traits. Given the very different appearance and time to maturity of the parents, it is not surprising that RIL was significant, but the analyses also show a significant interaction between Photoperiod and RIL. This suggests that the different RILs react differently to the different regimes, pointing to differences in sensitivity to photoperiod in the parents of the cross. Significant differences between trials for Photoperiod for all four traits indicates that the length of the day:night cycle affects flowering time and morphology irrespective of the amount of light received. However, this effect, while significant for all traits, explained most variation for flowering time, but only some for branching, and relatively little for biomass or height. Thus, biomass and height appear most affected by genotype and by amount of light received rather than by the day:night duration, pointing to the rate of carbon gain through photosynthesis as their main controlling factor.

The QTL analyses suggest a number of shared QTL regions along with multiple regions found only in individual trials that control flowering time, architecture, and biomass. Not surprisingly, QTL for flowering time, height, and biomass overlap in several regions, most notably on chromosome IV. There is less overlap with branching, reinforcing the conclusions of the ANOVA and correlation analyses that branching is controlled by a set of factors that are partially distinct from those for the other traits.

The major shared QTL region on chromosome IV has been shown to contain a number of genes involved in the photoperiod signaling pathway leading to flowering, including the Setaria orthologs of *HD1* (*CONSTANS*) and several copies of *FT*

(Mauro-Herrera et al. 2013). There is both a joint main QTL and a GxE QTL for branching in the same region, which also overlaps with a QTL for flowering time in a greenhouse trial and a QTL for branching in a field trial (Mauro-Herrera and Doust 2016). If one examines the QTL region of chromosome IV closely, it is apparent that the maximum LOD position for the QTL of the 8 and 12 h trials for flowering time and height are slightly offset from that for the 16 h trial, suggesting that the regulation of these traits in the long day 16 h trial is different from the short day 8 and 12 h trials. Evidence for this from the QTL analyses would be if it appeared more likely for QTL for the 8 and 12 h trials to group together than with QTL for the 16 h trial, but in fact it appears equally likely for 8 and 12 h trials to overlap as it is for 12 and 16 h trials to overlap. However, over all four traits there is only one example of QTL for the 8 h trial overlapping with the 16 h trial to the exclusion of the 12 h trial, which implies that the groupings of 8 and 12 h or 12 and 16 h QTL are nonrandom, suggesting a differentiation between shorter and longer photoperiods. Thus, the QTL analyses do give some support for separate short- and long-day responses in Setaria.

There is one other QTL region, apart from that on chromosome IV, where all three trials and the joint analysis have overlapping QTL. That is for height on chromosome IX, in the same region as the repressor of branching gene *teosinte branched1* (*tb1*) (Fig. 12.2). This was also found in both greenhouse and field trials. While it is possible that *tb1* is itself affecting height by repressing branch elongation, it is also possible that other genes in this region are involved.

There are a number of previously published QTL from greenhouse and field trials that overlap with QTLs found in the photoperiod growth chamber trials, but these do not appear to overlap any more frequently than QTL between the photoperiod trials. It is not surprising that QTL patterns differ between growth chamber, greenhouse, and field, as such patterns are well known in literature from other model systems, such as the difference in Arabidopsis mapping populations grown in greenhouse and field environments (Brachi et al. 2010; Malmberg et al. 2005). Those QTL regions that are constant between such varied environments, such as on chromosome IV for flowering time and V and IX for height and biomass should be investigated further for genes that differ between the parents and control these traits. There is less overlap of QTL for branching between environments indicating that this trait has strong and significant GxE interactions that govern the expression of the phenotype.

The QTL analyses are not sufficiently detailed to infer whether the different photoperiod regimes invoke different genetic pathways, in the manner of the differences between *HD3a* and *RFT1* expression under short and long days in rice. However, it is striking that the QTL intervals cover several of the major genes involved in the CONSTANS/*HD1* pathway but neither *EHD1* nor *GHD7*. It would be inappropriate to read too much into these analyses, but they suggest that further qRT-PCR analyses of plants at the floral transition should be undertaken to search for the participation of the *EHD1*/*GHD7* pathway in the regulation of photoperiod changes. The QTL analyses did not pick up significant differences between parental alleles at the *ZCN8* locus on chromosome III although our unpublished results do show that it is up-regulated at flowering. However, genome searches reveal that

several of the FT co-orthologs are present in QTL regions IV, VII, and VIII, making it possible that the parents differ in expression of the FT homologs but not the *ZCN8* homolog.

This study has uncovered interesting variation in the genetic regulation of flowering time and architectural traits and laid the stage for more intensive analyses. It has also shown that Setaria is variable in its architecture when grown under different environments (see also Chap. 10), suggesting that close attention needs to be paid to environmental conditions in order to understand phenotypic variation. The overlap between QTL identified in this study and in previous studies with genes in the CONSTANS photoperiod pathway suggests that variation in this pathway explains a significant proportion of the differences in flowering time seen between the two parents of the cross, as well as much of the variation in height and biomass. While not conclusive, the evidence presented here suggests that Setaria is a facultative short-day plant and that there may be differences in genetic regulation between short- and long-day photoperiod regimes. QTL for branching overlapped less frequently than those for height and biomass between trials and between this study and previous work, emphasizing the large environmental component to control of branching, and the weak relationship between branching and other architectural traits such as height and biomass gain. The insights gained in this study could not easily have been achieved in larger $C_4$ grasses such as maize or sorghum, and was only possible due to the small size, rapid life cycle and ease of growth of the Setaria system.

**Acknowledgements** I would like to thank Jessica Stromski for phenotyping and plant care and Margarita Mauro-Herrera for genetic analyses and fruitful discussions.

## References

Basten CJ, Weir BS, Zeng ZB, editors. Zmap-a QTL cartographer. 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software; 1994 August 7–12, Guelph, Ontario, Canada: Organizing Committee.

Basten CJ, Weir BS, Zeng ZB. QTL Cartographer Version 1.16 (1.16 ed.). Raleigh, NC: North Carolina State University; 2002.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30(6):555–61.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, et al. Linkage and association mapping of Arabidopsis thaliana flowering time in nature. PLoS Genet. 2010;6(5):e1000940.

Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. Genetics. 1994;138(3):963–71.

Doerge RW, Churchill GA. Permutation tests for multiple loci affecting a quantitative character. Genetics. 1996;142(1):285–94.

Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA. Genetic control of branching in foxtail millet. Proc Natl Acad Sci U S A. 2004;101(24):9045–50.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL. Foxtail millet: a sequence-driven grass model system. Plant Physiol. 2009;149(1):137–41.

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature. 2003;422(6933):719–22.

Higgins JA, Bailey PC, Laurie DA. Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses. PLoS One. 2010;5(4):e10065.
Hung HY, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, et al. ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc Natl Acad Sci U S A. 2012;109(28):E1913–21.
Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K. Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. Genes Dev. 2002;16(15):2006–20.
Jansen RC, Stam P. High-resolution of quantitative traits into multiple loci via interval mapping. Genetics. 1994;136(4):1447–55.
Jia GQ, Huang XH, Zhi H, Zhao Y, Zhao Q, Li WJ, et al. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nat Genet. 2013;45(8):957–61.
Komiya R, Yokoi S, Shimamoto K. A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. Development. 2009;136(20):3443–50.
Lazakis CM, Coneva V, Colasanti J. ZCN8 encodes a potential orthologue of Arabidopsis FT florigen that integrates both endogenous and photoperiod flowering signals in maize. J Exp Bot. 2011;62(14):4833–42.
Li PH, Brutnell TP. Setaria viridis and Setaria italica, model genetic systems for the Panicoid grasses. J Exp Bot. 2011;62(9):3031–7.
Malmberg RL, Held S, Waits A, Mauricio R. Epistasis for fitness-related quantitative traits in Arabidopsis thaliana grown in the field and in the greenhouse. Genetics. 2005;171(4):2013–27.
Mauro-Herrera M, Doust AN. Development and genetic control of plant architecture and biomass in the Panicoid grass, Setaria. PLoS One. 2016;11(3):e0151346. doi:10.1371/journal.pone.0151346.
Mauro-Herrera M, Wang XW, Barbier H, Brutnell TP, Devos KM, Doust AN. Genetic control and comparative genomic analysis of flowering time in Setaria (Poaceae). G3. 2013;3(2):283–95.
Meng X, Muszynski MG, Danilevskaya ON. The FT-like ZCN8 gene functions as a floral activator and is involved in photoperiod sensitivity in maize. Plant Cell. 2011;23(3):942–60.
Murphy RL, Klein RR, Morishige DT, Brady JA, Rooney WL, Miller FR, et al. Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. Proc Natl Acad Sci U S A. 2011;108(39):16469–74.
Murphy RL, Morishige DT, Brady JA, Rooney WL, Yang SS, Klein PE, et al. Ghd7 (Ma(6)) represses Sorghum flowering in long days: Ghd7 alleles enhance biomass accumulation and grain production. Plant Genome. 2014;7(2):1–10.
Saha P, Blumwald E. Spike dip transformation of Setaria viridis. Plant J. 2016;86:89–101.
Song YH, Ito S, Imaizumi T. Similarities in the circadian clock and photoperiodism in plants. Curr Opin Plant Biol. 2010;13(5):594–603.
Van Eck J, Swartwood K. *Setaria viridis*. In: Wang K, editor. *Agrobacterium* protocols. 2. New York: Springer; 2015. p. 57–67.
Vaughan DA, Lu BR, Tomooka N. The evolving story of rice evolution. Plant Sci. 2008;174(4):394–408.
Wolabu TW, Zhang F, Niu LF, Kalve S, Bhatnagar-Mathur P, Muszynski MG, et al. Three FLOWERING LOCUS T-like genes function as potential florigens and mediate photoperiod response in sorghum. New Phytol. 2016;210(3):946–59.
Yang SS, Murphy RL, Morishige DT, Klein PE, Rooney WL, Mullet JE. Sorghum phytochrome B inhibits flowering in long days by activating expression of SbPRR37 and SbGHD7, repressors of SbEHD1, SbCN8 and SbCN12. PLoS One. 2014;9(8):e105352.
Zeng ZB. Precision mapping of quantitative trait loci. Genetics. 1994;136(4):1457–68.

# Chapter 13
# Cell Wall Development in an Elongating Internode of *Setaria*

**Anthony P. Martin, Christopher W. Brown, Duc Q. Nguyen, William M. Palmer, Robert T. Furbank, Caitlin S. Byrt, Christopher J. Lambrides, and Christopher P.L. Grof**

**Abstract** Although *Setaria* has been proposed as a model to investigate $C_4$ photosynthesis, it may also be considered a suitable representative for biofuel feedstock species that are predominantly closely related panicoid grasses. In order to extend our understanding of the fundamental molecular and physiological mechanisms underpinning cell wall deposition as they occur during plant development, we have investigated an elongating stem internode of *S. viridis*. The chosen internode progressed from an active meristem and region of cell expansion at the base of the internode towards maturing fully expanded cells at the top of the internode. Along this developmental gradient, RNAseq of the mRNA fraction of the transcriptome was undertaken. A holistic understanding of the synthesis, composition and structure of the cell wall and the molecular mechanisms that signal the transition from primary to secondary cell wall synthesis will be integral to engineering crops with a structure that lends itself to more efficient deconstruction.

**Keywords** Cell wall • Stem • Internode • Transcriptome • Panicoid grass • Setaria

A.P. Martin • C.W. Brown • D.Q. Nguyen • W.M. Palmer • C.P.L. Grof (✉)
University of Newcastle, Newcastle, NSW, Australia

R.T. Furbank
Australian Research Council Centre of Excellence for Translational Photosynthesis, Plant Science Division, Research School of Biology, The Australian National University, Acton, ACT 2601, Australia

C.S. Byrt
School of Agriculture, Food and Wine, Waite Research Institute, University of Adelaide, Urrbrae, 5064, Australia

C.J. Lambrides
The University of Queensland, School of Agriculture and Food Sciences, Qld, 4072, Australia
e-mail: chris.grof@newcastle.edu.au



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_13

## 13.1 Introduction

The study of photoassimilate partitioning in plants has been a fundamental research area over many decades. The major focus of these studies has been to understand the allocation of carbon to some economic component of the plant, for example, grain in cereal crops, starch in root and tuber crops and sucrose in 'sweet' crops. The goal has been to increase harvest index, the proportion of the economic component relative to the total plant biomass. These studies are ongoing, but in recent years researchers have directed greater attention to other questions of carbon allocation. The economic importance of biomass has changed as plant cell walls are increasingly being used as a raw material for biofuel and biochemical production. This has influenced researchers to consider how plant cell walls are synthesised and what carbon resources are tied up in their production. It is not clear how much carbon could be reallocated from the synthesis of plant cell walls to other parts of the plant. To date, the majority of research on plant cell walls has focused on grain crops such as rice, wheat and maize, where the influence of grain cell wall composition is important for human nutrition (Collins et al. 2010). There is less information about the biology of plant cell walls in vegetative tissues such as the developing culm.

Herein, the genetic control of plant cell wall development in the culm is briefly reviewed and preliminary research investigating internode elongation of *Setaria*, a new model grass belonging to the Panicoideae, is introduced. An analysis of cell wall components of more than 200 *Setaria* germplasm lines has also been undertaken and is likely to provide the platform for more detailed investigation of plant cell walls in the developing culm of $C_4$ grasses. *Setaria* possesses the efficient $C_4$ pathway of photosynthesis and is therefore primed for high productivity under conditions of elevated light and temperature, similar to closely related species such as sugarcane (*Saccharum* spp.), Sorghum (*Sorghum bicolor* (L.) Moench), switchgrass (*Panicum virgatum* L.) and Miscanthus spp. Hence, in addition to its role as a model for $C_4$ photosynthesis (Brutnell et al. 2010), *Setaria* can also serve as a model system for photoassimilate partitioning in the developing culm.

## 13.2 Partitioning of Photoassimilate

Sucrose, the product of photosynthesis, is transported over long distances through the sieve elements of the phloem within the vascular system. Having reached growing sink organs, sucrose is hydrolysed by the enzyme sucrose synthase (SuSy) to produce the key molecule, UDP-glucose. Photoassimilate partitioning viewed from the position of the metabolite UDP-glucose can be considered in terms of three principal demands (Fig. 13.1). Demand 1 is carbon directed towards the structural components of cellulose and callose, catalysed by the enzymes cellulose and callose synthase, respectively, and non-cellulosic polysaccharides/pectin polymers catalysed by UDP-glucose dehydrogenase. To fulfil Demand 1 and complete the

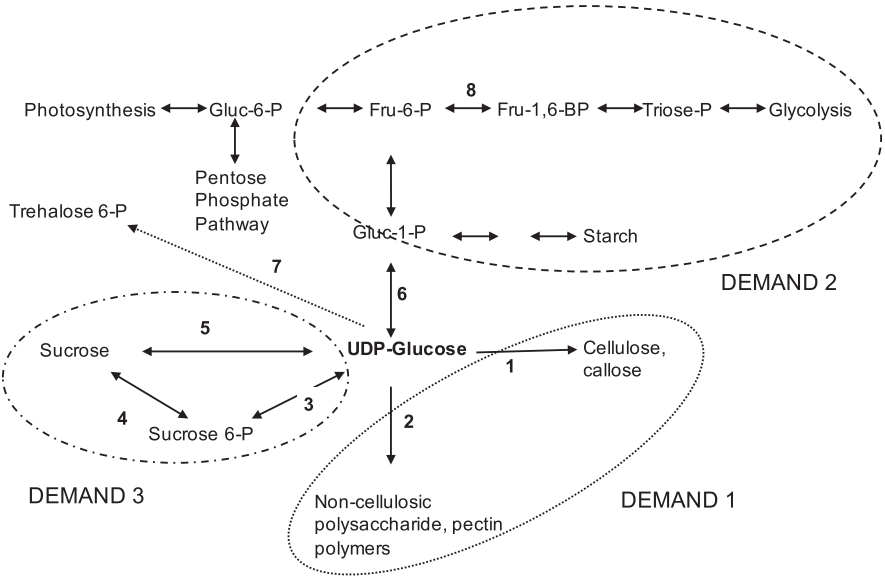


**Fig. 13.1** Photoassimilate partitioning in the growing sink organ, the culm, of *Setaria* and other monocot grasses. Photoassimilate partitioning, viewed from the central position occupied by the key metabolite UDP-glucose, considered in terms of three principal demands. Demand 1 is carbon directed towards structural components including cellulose/callose catalysed by cellulose/callose synthase (**1**), and non-cellulosic polysaccharide catalysed by UDP-glucose dehydrogenase (**2**). Demand 2 is carbon directed towards respiration, protein, lipid and starch synthesis, the first step being catalysed by UDP-glucose pyrophosphorylase (**6**). Demand 3 is carbon directed towards sucrose and catalysed primarily by the enzymes sucrose-phosphate synthase (**3**), sucrose-phosphate phosphatase (**4**) and sucrose synthase (**5**). Trehalose-phosphate synthase catalyses the synthesis of the sensor metabolite trehalose 6-phosphate (T6P) (**7**); Pyrophosphate:fructose 6-phosphate 1-phosphotransferase catalyses the reversible conversion of fructose 6-phosphate (Fru 6-P) and pyrophosphate (PPi) to fructose 1,6-bisphosphate (Fru 1,6-BP) and inorganic phosphate (Pi) (**8**)

synthesis of the growing cell walls, UDP-glucose can be interconverted to a range of five and six carbon nucleotide sugars through a series of complex enzymatic pathways (Bar-Peled and O'Neill 2011). Demand 2 is carbon directed towards respiration (glycolysis), lipid and starch biosynthesis, with the first step being catalysed by UDP-glucose pyrophosphorylase. Demand 3 is carbon directed towards sucrose and catalysed primarily by sucrose phosphate synthase (SPS), sucrose phosphate phosphatase (SPP) and SuSy. Trehalose phosphate synthase (TPS) catalyses the synthesis of the key sensor metabolite Trehalose 6-phosphate (T6P); T6P is also reliant on UDP-glucose as a substrate; however, the concentration of T6P is small and hence the demand upon UDP-glucose is minor. Pyrophosphate: fructose 6-phosphate 1-phosphotransferase catalyses the reversible conversion of fructose 6-phosphate (Fru 6-P) and pyrophosphate (PPi) to fructose 1, 6-bisphosphate (Fru 1,6-BP) and inorganic phosphate (Pi).

## 13.3 Cell Wall Development in an Elongating Internode

The culm, the central axis of the grass shoot, is made up of multiple repeating phytomeric units, comprising a node and an internode, where each node bears a leaf. Within each growing internode, the intercalary meristem resides at the base of the internode, immediately above the node below. Acropetally above the meristematic region, cells undergo expansion, maturing towards the top of the internode. Primary cell walls are synthesised whilst the cell is growing, whereas the secondary cell walls are produced as cell expansion ceases. The secondary cell wall is deposited inside the primary cell wall and characteristically is made up of three distinct layers (S1–S3) although in some fibre cells such as those of bamboo many more layers have been reported (Parameswaran and Liese 1976).

The plant body is made up of more than 30 different cell types, each of which is likely to vary in cell wall composition and structure (de Oliveira Buanafina and Cosgrove 2013). The complex nature of this extracellular layer is reflected in the number of genes, more than a thousand, that modulate cell wall construction and metabolism in maize (Penning et al. 2009). In broad terms, the cell wall is composed of cellulose intertwined in a complex matrix with non-cellulosic polysaccharides and pectin. The cellulose microfibrils are synthesised by a large protein complex embedded in the plasma membrane whilst the production of many non-cellulosic polysaccharides involves partial synthesis in the Golgi apparatus and delivery in secretory vessels to the surface of the cell for assembly. Non-cellulosic polysaccharides interact with cellulose to form the strong resilient structure that is the cell wall (Carpita 2012). In grasses, approximately 85 % of the primary cell wall is composed of (1,4)-β-glucans (cellulose), and non-cellulosic polysaccharides: arabinoxylan (AX), glucurono(arabino)xylan (GAX), xyloglucan (XyG) and (1,3;1,4)-β-glucans (mixed linkage glucans; MLG). The remaining components are small amounts of pectin, structural proteins, unpolymerised phenolics and silica. The secondary cell wall is structurally rigid and comprises cellulose, AX, GAX, silica and heavy deposits of the phenolic polymer lignin, which together make up approximately 95 % of secondary cell wall dry weight (Vogel 2008). There are also small amounts of pectin and unpolymerised phenolics such as ferulic and ρ-coumaric acid (see Table 13.1; adapted from (Vogel 2008)).

### *13.3.1 Cellulose*

Cellulose, the most abundant polymer on earth, is an unbranched, β-1,4 linked chain of D-glucose, successively inverted 180° and ranging in size from 2000 to 25,000 glucose residues (de Oliveira Buanafina and Cosgrove 2013; Richmond 2000; Taylor 2008). The repeating unit in a single cellulose chain is the dimer, cellobiose. The polymerisation of cellulose is catalysed by cellulose synthase (CesA) enzymes which belong to the Glycosyltransferase 2 superfamily (Campbell et al. 1997).

**Table 13.1** The approximate composition of primary and secondary cell walls in typical eudicots and grasses expressed as percentage of dry weight (% dry wt.)

| | Primary wall (% dry wt.) | | Secondary wall (% dry wt.) | |
|---|---|---|---|---|
| Components | Grass | Eudicot | Grass | Eudicot |
| **Cellulose** | 20–30 | 15–30 | 35–45 | 45–50 |
| **Non-cellulosic polysaccharide** | | | | |
| XyG | 1–5 | 20–25 | Minor | Minor |
| Xylans | 20–40 | 5 | 40–50 | 20–30 |
| MLG | 10–30 | Absent | Minor | Absent |
| Mannans and Glucomannans | Minor | 5–10 | Minor | 3–5 |
| **Pectins** | 5 | 20–35 | 0.1 | 0.1 |
| **Structural proteins** | 1 | 10 | Minor | Minor |
| **Phenolics** | | | | |
| Ferulic and $\rho$-coumaric acids | 1–5 | Minor | 0.5–1.5 | Minor[a] |
| Lignin | Minor | Minor | 20 | 7–10 |
| **Silica** | | | 5–15 | Variable |

After Vogel (2008)
[a]Except order Caryophyllales

Cellulose synthase-like (Csl) proteins encoded by the *Csl* genes also belong to this superfamily (Richmond and Somerville 2000).

Immunogold labelling of CesA enzymes has revealed that they are embedded in the plasma membrane and form a particle rosette consisting of six rosette subunits (Kimura et al. 1999). Each subunit is composed of six CesA enzymes which may be catalytically active as single polypeptide or dimer subunits (Emons and Mulder 1998; Carpita 2011). The eight highly conserved transmembrane domains of the proteins encoded by the CesA gene family defines the barrel shape of the rosette and the cytosolic orientation of the enzyme active site such that glucan chains are produced and secreted from the centre of the barrel into the extracellular space (Taylor 2008; Cosgrove 2005). This cellulose synthase enzyme complex (CSC) produces multiple glucan chains that weave and bind together spontaneously via hydrogen bonds to produce cellulose microfibrils, made up of between 12 and 36 interwoven glucan chains (Niimura et al. 2010; Fernandes et al. 2011; McFarlane et al. 2014). Recent NMR data and computational simulations of cellulose molecular dynamics indicate that 18- and 24-chain models are consistent with scattering and diffraction data; and spectroscopic measurements of Arabidopsis cellulose microfibrils indicated that the cellulose synthase complex is a hexamer of equimolar stoichiometry that synthesises an 18-glucan chain microfibril (McFarlane et al. 2014; Hill et al. 2014; Oehme et al. 2015). The cellulose synthase enzyme complex uses UDP-glucose as the substrate and the growing glucan chain is guided through the plasma membrane by cortical microtubules which ensure organised orientation of the cellulose microfibrils. This facilitates anisotropic cell wall expansion perpendicular to cellulose microfibril orientation and maximises strength in secondary cell walls (Mutwil et al. 2008).

A complement of ten genes encoding CesA has been identified in *A. thaliana* (Richmond and Somerville 2000; Doblin et al. 2010), ten in rice (Tanaka et al. 2003), 12 in sorghum (Paterson et al. 2009) and 12–14 in maize (Paterson et al. 2009; Appenzeller et al. 2004). Mutant analysis has demonstrated that at least three non-redundant CesAs form a functional cellulose synthase enzyme complex (CSC) and in rice, OsCesA 4, 7 and 9 are required for secondary cell wall synthesis (Tanaka et al. 2003). Expression analysis in maize has suggested that ZmCesA1 to 9 are involved in primary cell wall synthesis whilst ZmCesA10 to 12 are involved in secondary cell wall cellulose deposition (Appenzeller et al. 2004; Zhang et al. 2014). In *A. thaliana*, CesAs have been shown to be co-expressed in a range of tissues and demonstrate protein–protein interactions.

A number of other proteins have been proposed to play a role in cellulose synthesis, at least in Arabidopsis. Within the plasma membrane, CSI1 contributes to the mediation of CSC-microtubule alignment and in conjunction with KORRIGAN, a putative β-1,4 glucanase, may affect the motility of the CSC (McFarlane et al. 2014). Extracellular proteins, such as COBRA and CTL1, glycosylphosphatidylinositol (GPI)-anchored proteins, may also impact upon the velocity of the glucan chain production driven by the CSC (McFarlane et al. 2014). However, the precise mechanism of action of many of these ancillary proteins remains elusive. Mutations affecting primary cell wall cellulose synthesis tend to result in swollen cells, most likely due to isotropic cell expansion and the inability to regulate cell turgor in an expanding cell (Arioli et al. 1998). In the grasses barley, maize and rice, mutations in secondary cell wall cellulose synthesis genes consistently result in a brittle stem phenotype (Tanaka et al. 2003; Aohara et al. 2009; Sindhu et al. 2007; Kokubo et al. 1991), whereas in *A. thaliana* disruption of secondary cell wall cellulose synthesis resulted in collapsed xylem elements (Doblin et al. 2010).

### *13.3.2 Non-Cellulosic Polysaccharides*

The major non-cellulosic polysaccharides found in grasses, MLG, AX and XyG, are polymers of β-(1,4) linked sugars, often with a glucan backbone, making them structurally similar to cellulose. Non-cellulosic polysaccharides are often composed of a mixture of six and five carbon sugars and their backbone is often highly substituted with monosaccharide or disaccharide side chains (Doblin et al. 2010). In Type II commelinoid grasses, as for plants possessing Type I cell wall composition such as Arabidopsis, the widely accepted model describes non-cellulosic polysaccharides coating cellulose microfibrils, linking them together via hydrogen bonds and forming the cellulose-non-cellulosic polysaccharide network that confers mechanical strength to a cell wall (Park and Cosgrove 2015). For Type I cell walls, the microfibrils are tethered with xyloglucans and embedded in various pectins, and Type II cell walls are low in pectin.

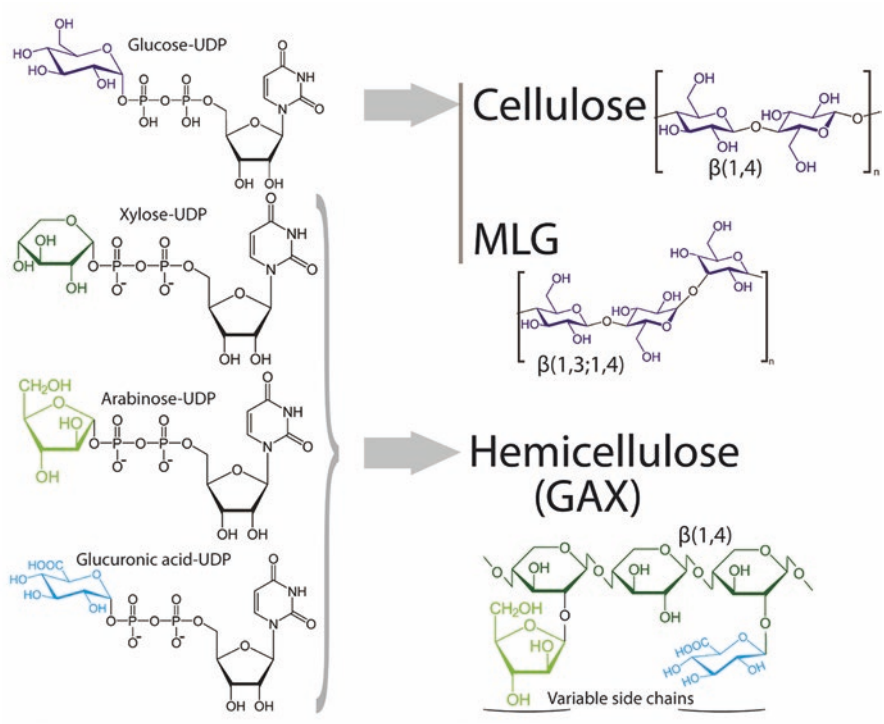


**Fig. 13.2** Major nucleotide sugar substrates required to build the cell wall polysaccharides of grasses (Demand 1). MLG; β(1,3;1,4)-glucans or mixed linkage glucans, G(AX); glucurono(arabino) xylan. Six carbon monomers are coloured in shades of *blue* and five carbon monomers are coloured in shades of *green* whilst the nucleotide functional groups are *black*

#### 13.3.2.1 Xylans

Glucurono(arabino)xylan (GAX) and arabinoxylan (AX), the predominant non-cellulosic polysaccharide in grass cell walls (Table 13.1; (Vogel 2008)), are polymers of β-1,4 linked D-xylose with variable non-repeating arabinose and glucuronic acid side chains for GAX and arabinose substitutions for AX (Fig. 13.2; (Faik 2010)). The synthesis of xylans occurs in the Golgi using UDP-xylose as the substrate for the xylan backbone (York and O'Neill 2008). The GAX xylose backbone is highly substituted with arabinose and glucuronic acid side chains during synthesis in the Golgi; however, these side chains are removed upon secretion to the wall and ferulic acid is often linked to the remaining arabinose side chains (Vogel 2008). This mode of synthesis has implications for cell wall molecular architecture. The arabinose and xylose side chains are usually attached at the O-2 and O-3 position (Vogel 2008), which blocks efficient hydrogen bonding to cellulose microfibrils and other GAX (Carpita 1983; McCann and Carpita 2008). Regulation and modulation

of side chain substitution may play a pivotal role in cell wall expansion of type II primary cell walls.

Xylan biosynthesis requires initiation, elongation and termination of the xylan backbone and, in the case of GAX, addition of arabinose and glucuronic acid side chains (Doblin et al. 2010). In *A. thaliana, GUX1* and *GUX2* genes encoding glucuronyltransferases have been localised to the Golgi and when transcription is repressed resulted in reduced addition of side chains to the xylan backbone; however, no reduction in synthesis of the xylan backbone was observed (Mortimer et al. 2010). This observation indicates that xylan backbone synthesis occurs via an independent mechanism to side chain substitution. In grasses, a number of glycosyltransferase (GT) families have been implicated in both xylan backbone synthesis and side chain substitution of xylans (Doering et al. 2012). Based upon a broad-scale bioinformatics approach (Mitchell et al. 2007), the GT43 gene family encoding β-1,4-xylan synthases, have been implicated in xylan backbone synthesis whereas the GT47 gene family, encoding xylan α-1,2- or α-1,3-arabinosyl transferases are proposed to direct arabinose substitutions of the xylan backbone. The heterologous expression of wheat and rice genes belonging to the GT61 family, *TaXAT2*, *OsXAT2* and *OsXAT3* in Arabidopsis *GUX1/GUX2* double mutants produced arabinosylated xylan, definitively demonstrating gain-of-function. The GT75 gene family is proposed to encode a glucuronosyltransferase required for the side chain addition of glucuronic acid (Zeng et al. 2010).

#### 13.3.2.2 Mixed Linkage Glucans

Mixed linkage glucans are a polymer of glucose with mixed β-1,3 and β-1,4 glycosidic linkages without any side chains (Vogel 2008). The mixed linkages between glucose monomers result in a polymer very similar to cellulose but with regular bends along the chain. They are present in high proportions (10–30 % cell wall dry weight) in expanding primary cell walls of grasses with maximum levels occurring at the peak of cell expansion (Kim et al. 2000), suggesting mixed linkage glucans play a major role in this developmental process.

Since MLGs are mostly specific to grasses, less is known about the molecular mechanisms of their synthesis. Identification of a quantitative trait locus for MLG content in barley grain led to the discovery of six rice cellulose synthase-like F (*CslF*) genes in the syntenic region of the sequenced rice genome (Burton et al. 2006). These *CslF* genes were heterologously expressed in *A. thaliana*, which does not produce MLG naturally and small amounts of MLG were detected (Burton et al. 2006). In addition, the *CslF* family is one of two grass-specific *Csl* gene families supporting the argument that *CslF* genes are responsible for biosynthesis of MLG, which is specific to grasses. Similarly, the *CslH* gene family, also specific to grasses has been shown to produce small amounts of MLG when heterologously expressed in *A. thaliana* (Doblin et al. 2009). Expression studies suggest that *CslF* genes are required in expanding primary cell walls, whilst *CslH* genes are required in seed endosperms or secondary cell walls where MLG is used as a carbon energy store

(Doblin et al. 2010). Both gene families are thought to only introduce β-1,4 glycosidic bonds into MLGs and since heterologous expression in *A. thaliana* only produces minute quantities of MLG it is considered likely that another unidentified gene family is responsible for producing the β-1,3 linkages in the MLG polymer (Doblin et al. 2010). Although there is a substantial body of immunochemical and biochemical data that supports the synthesis of full length mixed linkage β-glucans in the Golgi apparatus (Carpita 2012), preparation of grass tissues with high pressure cryofixation to preserve cellular ultrastructure and antigenicity revealed the presence of CSLF6 in the plasma membrane and intracellular membranes (Wilson et al. 2015), challenging the dogma that all non-cellulosic polysaccharides are synthesised and assembled in the Golgi complex.

#### 13.3.2.3 Xyloglucan (XyG)

Xyloglucan (XyG), a major component of eudicot cell walls, is only present in minor amounts in grasses (Table 13.1). It is a polymer of β-1,4 linked D-glucose with xylose monomeric and polymeric (xylose, galactose, fucose) side chains at the O-6 position in grasses. Three xylosyl residues are substituted in repeating units of four glucose backbone monomers (Pauly et al. 2013). Following synthesis in the Golgi complex, directed by synthases encoded by the *CslC* gene clade, xyloglucan is transported to the cell membrane. The number of CslC enzymes contributing to biosynthesis and the mechanism for side chain addition remains largely undescribed in grasses (Park and Cosgrove 2015; Pauly et al. 2013; Scheller and Ulvskov 2010).

#### 13.3.2.4 Pectins

Pectin is a family of complex galacturonic acid-rich polysaccharides comprising ~35 and 5% of eudicot and grass cell walls, respectively. They consist of a galacturonan backbone that can be substituted at various positions with simple sugars including xylose, rhamnose and apiose, as well as complex side chains (Harholt et al. 2010). Four types of pectin polysaccharides have been described namely, homogalacturonan (HG), xylogalacturonan (XGA), apiogalacturonan, rhamnogalacturonan I (RGI) and rhamnogalacturonan II (RGII) with HG accounting for approximately 65% of total pectin in plants. Pectins, particularly HG and RGII, play a principal role in shaping and strengthening the cell wall by acting as adhesive agents in the middle lamella (Harholt et al. 2010).

### *13.3.3 Phenolics*

Phenolics, secondary metabolites which are both extremely abundant and widely distributed in the plant kingdom, play important functional roles in pigmentation, growth, reproduction and pathogen resistance (Cheynier et al. 2013; Lattanzio et al.

2006). They possess one or more aromatic rings bearing one or more hydroxyl branches and are derived from the Shikimate-Phenypropanoid pathway (Cheynier et al. 2013; Dai and Mumper 2010). In the cell walls of grasses, the significant phenolic compounds identified include the hydroxycinnamates, ferulic acid and ρ-coumaric acid, and the polyphenol lignin which is a complex polymer that varies in monomer composition between species and cell types.

#### 13.3.3.1 Lignin

Lignin is the third most abundant heteropolymer (approximately 20 % dry weight) in secondary cell walls of grasses, predominantly associated with vascular bundles. It is a complex aromatic polymer generated by irregular linkage of three main phenylalanine-derived monolignols, Hydroxyphenyl (H), Guaicyl (G) and Syringyl (S) (Li et al. 2008), with the resultant polymer being highly branched due to a variety of possible linkages (β-O-4, β-β, β-5) between monomers. The incorporation of lignin in plant cell walls through covalent binding to non-cellulosic polysaccharides provides rigidity, strength and a barrier against external physical forces and pathogen attack. Lignin also renders the xylem impermeable to water and solutes, thereby facilitating water and nutrient transport.

H lignin is uniquely present in grass cell walls with a typical grass cell wall containing ~35–49 % G units, 40–61 % S units and 4–15 % H units (Vogel 2008; Vanholme et al. 2010). The ratio of S:G lignin has been implicated in defining cell wall characteristics since G lignin is more highly branched than S lignin and therefore has greater opportunity to cross link cell wall polysaccharides. A lower S:G ratio therefore, typically results in a more rigid and less digestible cell wall (Abreu et al. 2009; Koutaniemi 2007).

The highly controlled process of lignification can be considered in terms of three principal stages (1) biosynthesis of monolignols in the cytosol; (2) transportation of these precursor monolignols across the plasma membrane; and (3) lignification of the monolignols by oxidative polymerisation (Liu 2012). Stage 1 occurs on the cytosolic side of the ER, where chloroplast-derived phenylalanine is de-aminated by the enzyme phenylalanine ammonia lyase (PAL) to form cinnamic acid (Li et al. 2008). Successive hydroxylation and methylation reactions form ρ-coumaroyl CoA, an important metabolite and branching point of monolignol or flavonoid biosynthesis pathways. At this point, monolignol biosynthesis also diverges to produce ρ-coumarate aldehyde and ρ-coumaroyl shikimic acid through the catalytic activity of hydroxycinnamoyl CoA reductase (CCR) and hydroxycinnamoyl CoA shikimate: quinate hydroxycinnamoyl transferase (HCT), respectively. The aldehyde is then converted into the ρ-coumaryl alcohol, catalysed by hydroxycinnamyl alcohol dehydrogenase (CAD) and subsequently converted into the H monomeric lignin subunit (Hao and Mohnen 2014). In the alternate arm of the pathway, ρ-coumaroyl shikimic acid undergoes serial reduction reactions by ρ-coumaroyl shikimate 3′-hydroxylase (C3H), ρ-hydroxycinnamoyl CoA shikimate:quinate ρ-hydroxycinnamoyl transferase; (HCT/CST), Caffeoyl CoA O-methyl transferase

(CCoAOMT) and (hydroxy)cinnamoyl CoA reductase (CCR) to produce coniferaldehyde. Synapaldehyde is produced by the catalytic activity of ferulic acid/coniferaldehyde/coniferyl alcohol 5-hydroxylase (F5H) and caffeic acid/5-hydroxyferulic acid O-methyl transferase (COMT) in series upon coniferaldehyde. Synapy and coniferyl alcohol, produced by the activity of (hydroxyl) cinnamyl alcohol dehydrogenase (CAD), are transported through the plasma membrane to the apoplasm and oxidised to produce the remaining two monomeric lignin subunits, Syringyl and Guaicyl lignin (Boerjan et al. 2003; Zhong and Ye 2015).

Three possible mechanisms have been proposed for the transport of the monolignols to the apoplasm for polymerisation. The hydrophobic monolignols may diffuse passively through the plasma membrane thereby preventing their toxic accumulation in cells (Boija and Johansson 1758). The widely accepted mechanism of vesicular secretion of the monolignols into the extracellular space by the Golgi complex has recently been challenged. TEM (transmission electron microscopy) autoradiographs of radiolabelled monolignols in cells undergoing lignification in Lodgepole pine (*Pinus contorta* Dougl. Ex Loud.) tracheids have allowed monolignol transport to be visualised, revealing that they are not translocated via the Golgi (Kaneda et al. 2008). An alternative mechanism of lignin subunit transport has been proposed, involving ATP-binding (ABC) transporters, initially based upon transcript profiling, which shows tight co-expression of some ABC transporters with monolignol biosynthesis genes (Ehlting et al. 2005). ABC transporters belong to a gene superfamily with 130 putative members reported in Arabidopsis (Kang et al. 2011) and more than 130 in rice (Saha et al. 2015). They are considered to be responsible for the transport of metabolites, signalling molecules lipids and proteins across cell membranes, with some members demonstrated to be phenolic/polyphenolic transporters (Sibout and Höfte 2012). The putative ABC transporter AtABC29, highly expressed in Arabidopsis roots and anthers, has been identified as a monolignol transporter by both microarray and experimental analysis (Alejandro et al. 2012). Mutant Arabidopsis knockout plants of AtABC29 demonstrate increased levels of ρ-coumaryl alcohol in the cytosol and reduced lignification (Alejandro et al. 2012). Six closely related ABC transport members that belong to the pleiotropic drug resistance (PDR) subfamily have been proposed as potential transporters responsible for trafficking of the two remaining monolignols through the plasma membrane (Sibout and Höfte 2012).

Once delivered to the apoplasm, the monolignols are primed for polymer assembly by dehydrogenation (reduction of the HCA-CoA esters) catalysed by several enzymes including peroxidases, laccases and polyphenol oxidases (Boerjan et al. 2003; Sibout and Höfte 2012). The monolignol radicals are randomly cross-coupled to one another or to a growing polymer complex via endwise coupling predominantly by β-O-4 linkages to form the three dimensional, branched, interlocking lignin networks (Vanholme et al. 2010; Hao and Mohnen 2014; Boerjan et al. 2003; Hatfield and Vermerris 2001). Based upon micro-autoradiography, lignification is initiated at $Ca^{2+}$-rich nucleation sites, as calcium is required for the polymerisation of coniferyl alcohol, where lignin is covalently linked with the non-cellulosic polysaccharide and pectin constituents of the primary cell wall (Hao and Mohnen 2014). During the early stages of cell wall lignification, H lignin is believed to be deposited

and functions as a scaffold to determine cell shape prior to deposition of G and S lignin (Nakashima et al. 2008). In the grasses sugarcane and rice, lignification begins in the cell walls of the protoxylem vessels, progresses to the middle lamellae of fibre cells, the secondary wall of metaxylem vessels and finally to the secondary wall of the fibres (He and Terashima 1991; de Oliveira Buanafina and Cosgrove 2013).

#### 13.3.3.2 Hydroxycinnamates

The hydroxycinnamates, ferulic acid and ρ-coumaric acid, are intermediates in the monolignol biosynthesis pathway and constitute up to 4 and 3 % of cell wall dry weight, respectively (Vogel 2008). A unique feature of grass cell walls is the covalent ester linkage of ferulic acid to arabinose side chains of arabinoxylan or glucoarabinoxylan (de Oliveira Buanafina 2009). The arabinoxylan linked ferulates may form dimers, trimers and tetramers that cross link adjacent arabinoxylans during cell wall deposition and lignification (Bunzel et al. 2006). Furthermore, ferulic acid is able to provide the bridge to form lignin-ferulate-polysaccharide complexes through ester-ether linkages (de Oliveira Buanafina and Cosgrove 2013; Jacquet et al. 1995) thereby filling the functional role of structural proteins in Type I cell walls by cross linking non-cellulosic polysaccharide, cellulose and lignin constituents (Vogel 2008). It has also been suggested that ferulic acids act as initiation sites for lignin polymerisation within the secondary cell wall (Ralph et al. 1995) and genetically engineering monolignol ferulate conjugates specifically into poplar xylem significantly increased cell wall digestibility (Wilkerson et al. 2014).

## 13.4 Growth of the Cell Wall

The widely accepted architectural model of primary cell walls of both eudicots and monocots consists of cellulose microfibrils coated in non-cellulosic polysaccharides and tethered together covalently by hydrogen bonds to create a strong matrix network. Growth of the cell involves irreversible expansion of the cell wall in conjunction with an influx of water into the cell. The growth process, culminating in the synthesis, secretion and intercalation of new moieties into the developing cell wall requires vacuolar enlargement and solute uptake to maintain osmotic potential, hence turgor pressure.

Cell growth is physically constrained by the cell wall and dogma dictates that enzymes act specifically upon the non-cellulosic polysaccharide tethers (MLGs in grasses) to loosen their links to the cellulose microfibrils, hence promoting slippage or ‘wall loosening’. However, plant endoglucanases and endotransglycosylase/hydrolases do not significantly influence cell wall relaxation primarily because the xyloglucans accessible to these enzymes are not ‘load bearing’ (Park and Cosgrove 2012). Furthermore, solid-state Nuclear Magnetic Resonance data generated from mung bean cell walls (Bootten et al. 2004), indicated that less than 8 % of cellulose surfaces are coated with xyloglucan.

Empirical evidence indicates that the widely accepted model of primary cell wall architecture needs to be revisited and the tethering role of xyloglucans reconsidered. A 'biomechanical hotspot' hypothesis has been proposed (Park and Cosgrove 2015), whereby limited cellulose-cellulose contact takes place at mechanical junctions possibly mediated by a xyloglucan monolayer binding the hydrophobic surfaces together. Cell wall extensibility may be promoted by expansin activity, through some undescribed mechanism, at these sites. Pectins are also proposed to play a significant role, as contact between pectins and the cellulose microfibrils is estimated to be greater than that of xyloglucan.

## 13.5 Molecular Characterisation of the Elongating *Setaria* Internode

One way to investigate the genes involved in internode development is to analyse the gene transcript levels. Although highly complex, cell walls are constructed from a limited number of components, predominantly cellulose microfibrils, non-cellulosic polysaccharides, lignin and pectin; and many of the genes involved in making these components are known. To begin unravelling this complexity in *Setaria*, gene expression in internodes of *Setaria* undergoing development from cellular division, to cellular expansion of the primary cell wall and culminating in secondary cell wall deposition have recently been analysed (Fig. 13.3; (Martin et al. 2016)). Similar strategies have been used for other grass species including rice (Hirano et al. 2013), maize (Zhang et al. 2014; Bosch et al. 2011) and sugarcane (Casu et al. 2007). Previous maize (Bosch et al. 2011) and sugarcane studies (Casu et al. 2007) undertook microarray transcriptomic analysis comparing entire elongating and fully elongated culm internodes. In previous studies of elongating internodes, differences in gene expression relative to mature fully elongated internodes have been observed; however, some changes are likely to be dampened by homogenisation of the different developmental regions within the internode during sample preparation. Recently, a more refined approach was taken with an elongating maize internode, where sections 10 cm in length were divided into 1 cm sections prior to detailed analysis (Zhang et al. 2014). By studying 1 cm sections, the authors observed that the expression of genes encoding key enzymes in the biosynthesis and modification of cellulose, non-cellulosic polysaccharide (predominantly glucuronoarabinoxylan) and lignin, peaked in the transitional region of internode development prior to completion of the deposition of secondary cell walls.

### *13.5.1* Setaria CesA *and* Csl *Genes*

Synthases involved in producing the cell wall 'backbone' are encoded by members of the *CesA/Csl* superfamily of genes (Carpita 2011). Cellulose is synthesised by members of the *CesA* clade; mannans and glucomannans by members of the *CslA* clade;

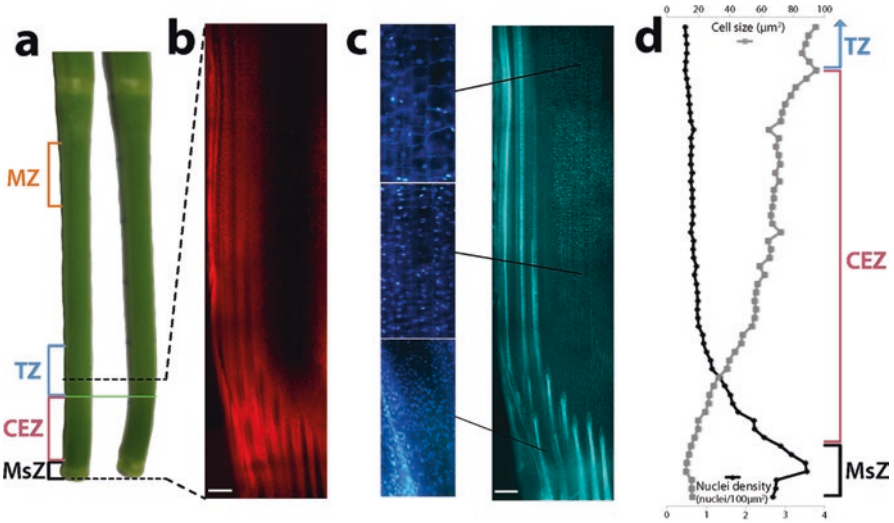


**Fig. 13.3** Regions of *Setaria* internode 5 of 7 (from base) selected for RNA isolation and sequencing. (**a**) Internode 5 harvested at the 'half head emergence' developmental stage with its leaf sheath stripped, alongside an equivalent internode where the lower, flexible zone has been bent to display the difference between the upper rigid and the lower more flexible zone. The *green* line indicates the interface between the flexible and rigid zones of the internode. The harvested meristematic (MsZ—*black*), cell expansion (CEZ—*pink*), transitional (TZ—*blue*) and the mature (MZ—*orange*) zones are indicated. (**b**) and (**c**), Longitudinal, vibratome cut, 50 μm thick section of the lower region of the internode (indicated by *dotted black lines*), stained with DAPI and viewed under UV illumination with (**b**) the *red* chlorophyll emissions isolated and (**c**) the *blue* DAPI emissions isolated with enlarged regions offset (*white* scale bars are 50 μm). (**d**) Nuclei density (*black circles*; nuclei/100 μm$^2$) and cell size (*grey squares*; μm$^2$) measured using image J were plotted at intervals along the lower region of the internode. This figure was adapted from Martin et al. (2016)

xyloglucans, by members of *CslC* clade, and mixed-linkage β-glucans, by members of the *CslF* and *CslH* clades (Carpita 2012). Ten of the 13 *Setaria CesA* genes identified here exhibited a greater than 1 $Log_2$ fold change in the region of cell elongation as compared to the meristematic region (Fig. 13.4). Most striking is the fourfold $Log_2$ change in three of the *CesA* genes (*CesA 4, 10 and 12*) in the transitional and maturing regions of the internode. Phylogenetically, these three genes are most similar to *CesA4 (*At5g44030*), 7* (At5g17420) and *8* (At4g18780) from Arabidopsis and *CesA10*, *11*, *12(13)* from maize, which have been implicated in secondary cell wall cellulose synthesis (Appenzeller et al. 2004). This sustained elevated expression is at odds with the report of ongoing but reduced expression of a broad cohort of genes in the maturation region of the maize internode (Zhang et al. 2014).

Genes belonging to the *CslA* clade were highly expressed in *Setaria* (with one exception, *CslA11*). *CslA* expression was particularly high in the meristematic and elongation zones of the internode (Fig. 13.5). In maize, the cell walls within the elongation zone were measured to contain 3 % mannan and then decreased acropetally in the internode (Zhang et al. 2014), matching the *Setaria CslA* gene expression reported here. The genes belonging to the *CslC* subgroup putatively encode the

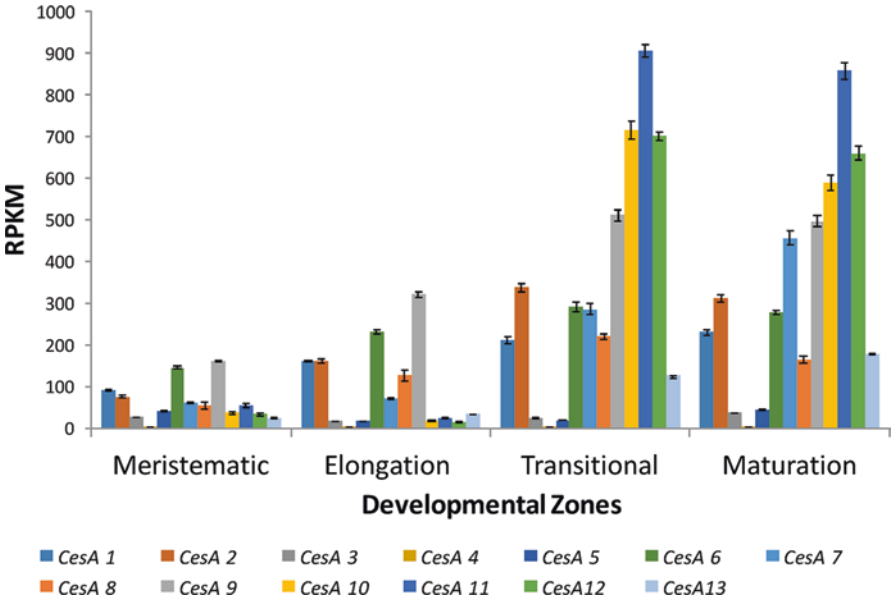


**Fig. 13.4** Expression profile of *CesA* genes in an elongating internode of *Setaria*. RNAseq reads per kilobase per million reads (RPKM) expression levels of all identified cellulose synthase A (*CesA*) transcripts in *S. viridis* showing maximal *CesA* expression during the transitional stage of internode development. An average of four biological replicates ± SE are displayed in the meristematic, cell elongation, transitional and mature zones of the internode. Data was obtained from Martin et al. (2016)

xyloglucan synthases and these exhibited an elliptical profile, with highest expression in the elongation zone and relatively lower expression in the transitional and maturing zones of the internode.

*CslF6,* a member of the *CslF* subgroup exhibited greater expression in the transitional zone relative to the meristematic zone, and this higher expression was also observed in the maturing region of the internode (Fig. 13.5). The expression of *CslH2* was highest in the elongation zone and close to zero in the transitional and maturation regions of the internode. The expression of *CslH1* increased with the maturity of the internode although the level of expression was extremely low.

### *13.5.2* Setaria *Heteroxylan Synthase Genes*

Xylan, represented in grasses as AX or GAX, makes up between 20 and 40 % (dry weight) of the primary cell wall and between 40 and 50 % of the secondary cell wall composition (Table 13.1; (Vogel 2008)). In maize, GAX makes up more than 30 % of the primary cell wall and close to 50 % in the secondary cell walls (Zhang et al. 2014). The glycosyltransferases contributing to the synthesis of heteroxylans are encoded by members belonging to a large superfamily of genes (Hansen et al. 2012). *Setaria* genes belonging in the GT family 47 (Fig. 13.6)

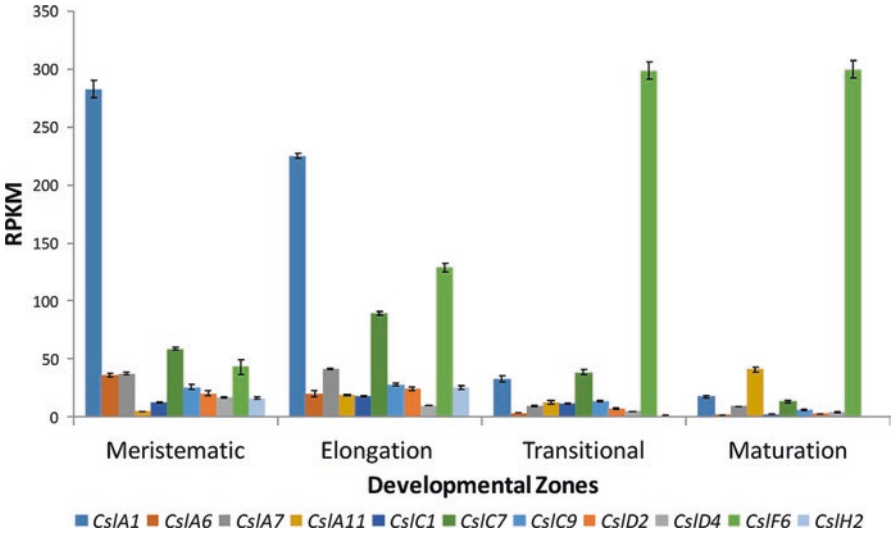


**Fig. 13.5** Expression profile of *Csl* genes in an elongating internode of *Setaria*. RNAseq reads per kilobase per million reads (RPKM) expression levels of all identified cellulose synthase-like (*Csl*) transcripts in *S. viridis* showing differential expression of *Csl* genes through internode development. An average of four biological replicates ± SE are displayed in the meristematic, cell elongation, transitional and mature zones of the internode. Data was obtained from Martin et al. (2016)

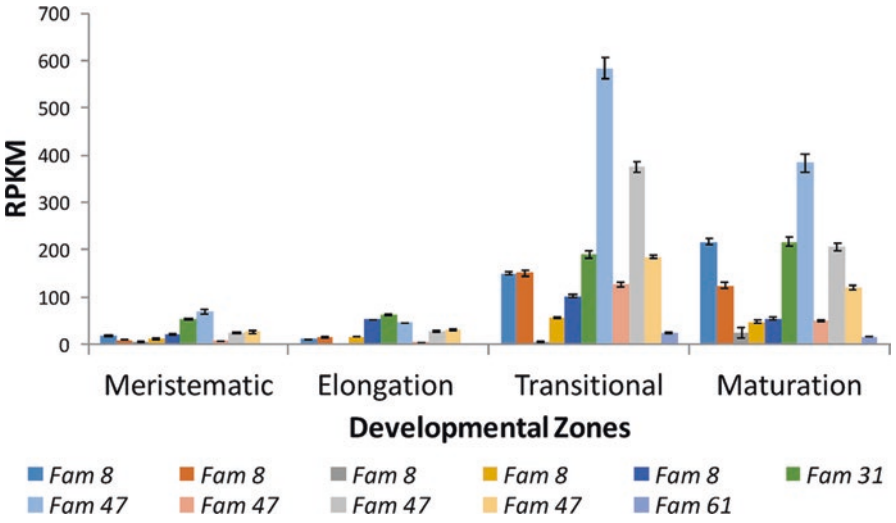


**Fig. 13.6** Expression profile of glycosyltransferase (GT) genes in an elongating internode of *Setaria*. RNAseq reads per kilobase per million reads (RPKM) expression levels of identified glycosyltransferase (*GT*) transcripts (Family (*Fam*) 8, 31, 47 and 61) in *S. viridis* showing maximal expression of *GT* genes in the transitional zone of the developing internode. An average of four biological replicates ± SE are displayed in the meristematic, cell elongation, transitional and mature zones of the internode. Data was obtained from Martin et al. (2016)

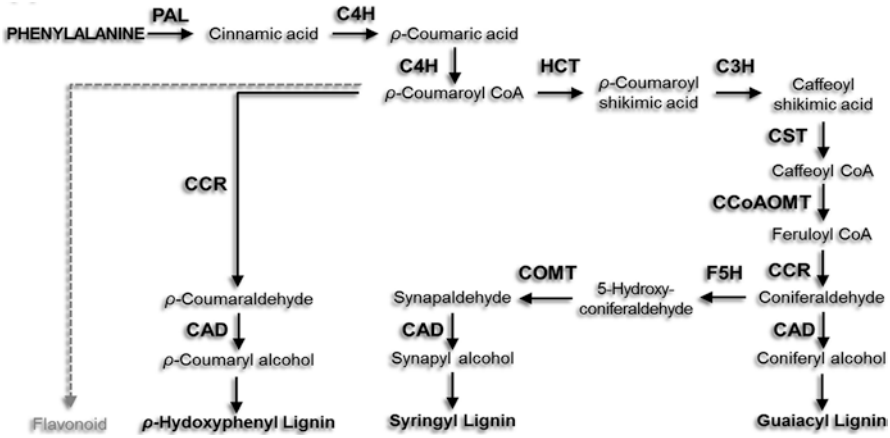


**Fig. 13.7** The lignin biosynthesis pathway. This model depicts the ten enzymatic steps leading to monolignol synthesis. The hydroxycinnamyl alcohols produced, Guaiacyl, ρ-Hydroxyphenyl and Syringyl polymerise to form lignin. The enzymes and abbreviations are: phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate-CoA ligase (4CL), hydroxycinnamoyl CoA:shikimate transferase (HCT), p-coumarate 3-hydroxylase (C3H), caffeoyl CoA O-methyltransferase (CCoAOMT), cinnamyl CoA reductase (CRR), ferulate 5-hydroxylase (F5H), caffeic acid *O*-methyltransferase (COMT) and cinnamyl alcohol dehydrogenase (CAD). Modified from Sattler and Funnell-Harris (2013)

exhibit high levels of expression and substantial $Log_2$ fold changes in the transitional and maturing regions of the internode, consistent with these glycosylases being involved in elongation of the xylan backbone. Similarly, a number of members of the GT family 8 increase strongly in expression in these same regions, consistent with their purported role in xylan backbone substitution (Rennie and Scheller 2014).

### *13.5.3* Setaria *Lignin Biosynthesis Genes*

The genes encoding the ten enzymes participating in catalytic steps of lignin biosynthesis (See Fig. 13.7) are well known (Vanholme et al. 2010). The most highly expressed representatives of each group in *Setaria* are illustrated in Fig. 13.8. All of the *Setaria* lignin biosynthesis genes analysed here are considered to be 'highly enriched in the internode' in that the $Log_2$ fold expression is many times higher than expression of the same gene in a 'whole plant' transcriptome analysis encompassing multiple tissues across three developmental stages of growth, namely seed germination, vegetative growth and reproduction (Xu et al. 2013).

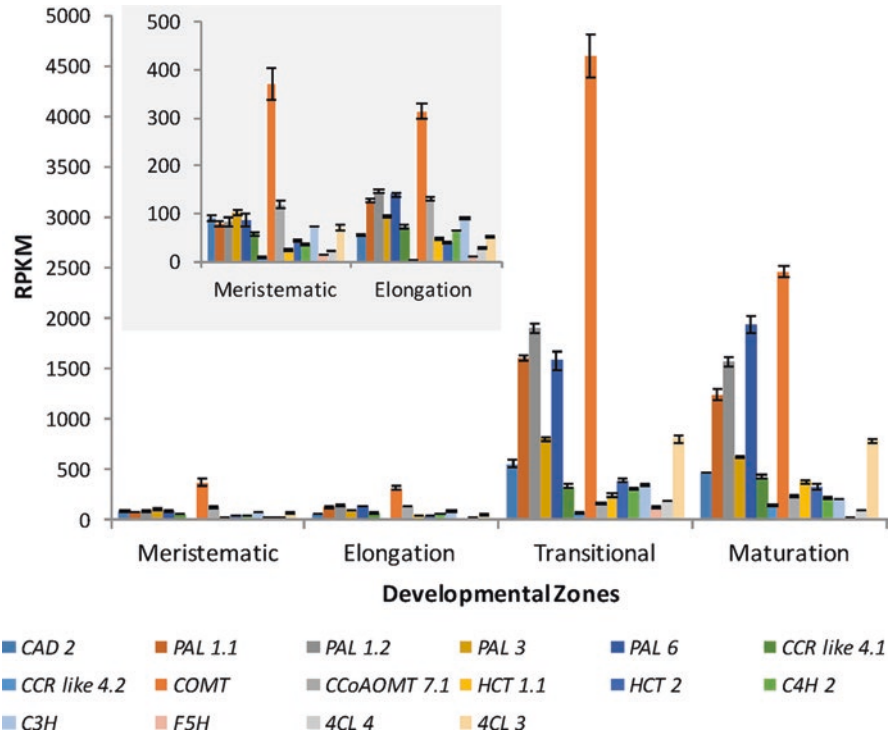


**Fig. 13.8** Expression profile of genes encoding enzymes involved in lignin biosynthesis in an elongating internode of *Setaria*. RNAseq reads per kilobase per million reads (RPKM) expression levels of identified lignin biosynthesis enzyme transcripts in *S. viridis* showing maximal expression in the transitional zone of the developing internode. Phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate-CoA ligase (4CL), hydroxycinnamoyl CoA:shikimate transferase (HCT), p-coumarate 3-hydroxylase (C3H), caffeoyl CoA O-methyltransferase (CCoAOMT), cinnamyl CoA reductase (CRR), ferulate 5-hydroxylase (F5H), caffeic acid O-methyltransferase (COMT) and cinnamyl alcohol dehydrogenase (CAD). An average of four biological replicates ± SE are displayed in the meristematic, cell elongation, transitional and mature zones of the internode. Data was obtained from Martin et al. (2016)

## 13.6 Transcript Profiling of Carbon Demands

An overview of the levels of cell wall gene transcripts and other genes associated with the partitioning of photoassimilate (see Fig. 13.1) within the context of a developing *S. viridis* internode was recently published (Martin et al. 2016).

As expected, within the developing internode, the 'meristematic zone' and 'cell expansion zone' showed gene expression related to metabolism/energy production, some aspects of cell wall synthesis, namely callose synthases, hemicellulose synthesis via GT families and AGPases implying starch accumulation. Plasmodesmatal proteins were active, which indirectly implied that the carbon source was predominantly sug-

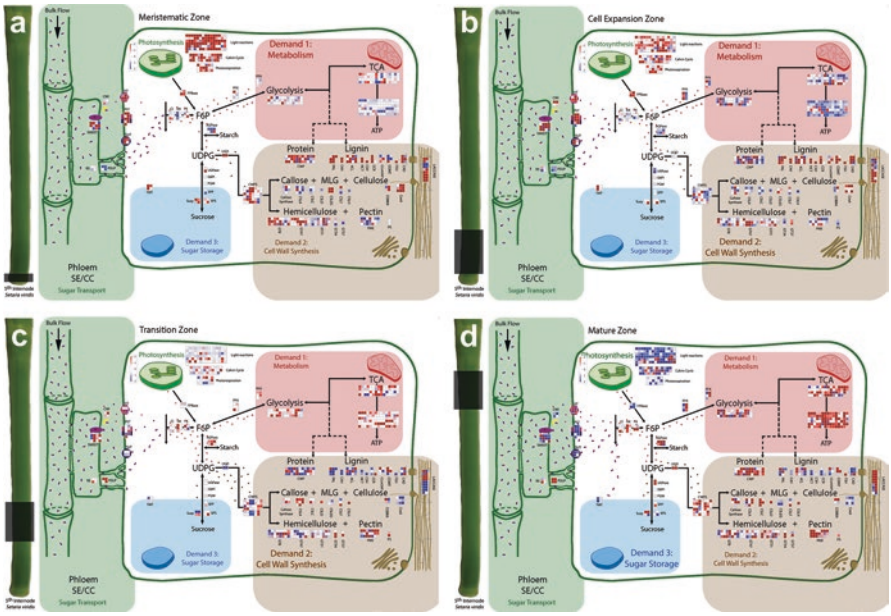


**Fig. 13.9** Schematic model of carbon demands during internode development based on transcriptomic analysis. A schematic pathway of carbon import into the developing *S. viridis* fifth internode, showing gene expression relating to the three major demands on carbon supply. Genes were assigned into categories (as labelled on the diagram) and their expression values were displayed as squared $\log_2$ fold change from the squared average of the four internode zones for each gene. Up-regulated genes are shown in *blue*, whilst down-regulated genes are in *red. MLG* mixed linkage glucans, *RS* raffinose synthase, *PDLP* plasmodesmata localised proteins, *SWEET* sugars will eventually be exported transporters, *CWI* cell wall invertase, *HT* hexose transporters, *PLT* polyol transporters, *SUT* sucrose transporters, *CI* cytosolic invertases, *HK* hexokinase, *FK* fructokinase, *FPBase* fructose-1,6-bisphosphatase, *F6P* fructose-6-phosphate, *PFK* phosphofructokinase, *AGPase* glucose-1-phosphate adenylyltransferase, *UGD* UDP-glucose 6-dehydrogenase, *UGPase* UDP-glucose phosphorylase, *G6PI* glucose-6-phosphate isomerase, *PGM* phosphoglucomutase, *SPP* sucrose phosphate phosphatase, *SPS* sucrose phosphate synthase, *Susy* sucrose synthase, *TMT* tonoplast membrane transporter, *CWPS* cell wall precursor synthesis, *CWP* cell wall protein, *PAL* phenylalanine ammonia-lyase,*C4H* cinnamate-4-hydroxylase, *4CL* 4-hydroxycinnamate CoA ligase, *HCT* hydroxycinnamoyl transferase, *C3H* coumarate 3-hydroxylase, *CCR* cinnamoyl-CoA reductase, *CCoAOMT* caffeoyl CoA 3-*O*-methyltransferase, *COMT* caffeic acid *O*-methyltransferase, *F5H* ferulate 5-hydroxylase, *CAD* cinnamyl alcohol dehydrogenase, *CSLAnH* cellulose synthase-like A-H, *CesA* cellulose synthase complex, *GT* glycosyltransferase, *PME* pectin methylesterase, *PS* pectin synthesis, *SE* sieve element, *CC* companion cell. *Blue* and *purple dots* are representative of sugar flow. (**a**) Meristematic Zone; (**b**) Cell Expansion Zone; (**c**) Transition Zone; (**d**) Mature Zone. Reproduced from Martin et al. (2016)

ars imported through symplasmic connections from sieve elements containing sugars delivered to the growing sink tissue by bulk flow (Fig. 13.9).

The major carbon demand in the 'transitional zone' of the internode shifted to cell wall synthesis, specifically lignin, cellulose, cell wall proteins and some GT families (hemicellulose synthesis). UDP-glucose 6-dehydrogenase (UGD) and cell

wall precursor synthesis (CWPS) proteins were highly expressed showing how carbon from UDP-glucose pools is directed into cell wall synthesis at this stage of internode development, where cell expansion ceases and extensive deposition of secondary cell walls occurs (Fig. 13.9).

In the 'mature zone' of the developing internode, the cell walls are now thickened secondary walls, and the internode can realise its capacity for sugar accumulation. This is indicated by the expression of sugar transporter genes from the polyol transporter (PLT), hexose transporter (HT), sucrose effluxer (SWEET) and sucrose transporter (SUT) families. Photosynthesis genes are also active in this mature region, indicating an additional carbon source that does not require long distance transport of sugars. Tonoplast monosaccharide transporters (TMTs) also show increased expression in this zone indicating their role in sugar storage within vacuoles in mature internodes (Fig. 13.9).

This recently published overview of carbon partitioning within a developing internode provides a valuable platform for gene discovery; not only identifying gene families that are active at different stages of internode development, but also allowing identification of the dominant isoforms from each family for genetic manipulation and further experimental study of internode development in a $C_4$ grass.

## 13.7 Germplasm to Support the Molecular Dissection of Plant Cell Wall Components in *Setaria*

We assembled 214 ecotypes of *Setaria* from around the world and evaluated them in a glasshouse experiment for agronomic traits and percent lignin. The plants were grown under controlled conditions 24 °C/20 °C day/night temperature in 2013. There was large variation for all traits measured (Table 13.2) including height, maturity (anthesis), number of internodes, biomass, stem width, tiller number, number of leaves and percent lignin. Not surprisingly, biomass and other growth

**Table 13.2** Mean, minimum and maximum values for several attributes measured in 214 *Setaria italica* germplasm lines grown in a controlled environment glasshouse at The University of Newcastle, NSW Australia in 2013

| Attribute | Mean | Minimum | Maximum |
|---|---|---|---|
| Height (mm) | 1276.0 | 110.0 | 2673.0 |
| Maturity (days to anthesis) | 69.2 | 30.0 | 201.0 |
| Number of internodes | 10.3 | 3.0 | 19.0 |
| Biomass (g dry weight) | 3.8 | 0.04 | 15.9 |
| Stem width (mm) | 10.7 | 3.0 | 19.0 |
| Number of tillers | 2.5 | 1.0 | 21.0 |
| Number of leaves | 22.3 | 5.0 | 72.0 |
| Percent lignin | 20.3 | 16.6 | 23.8 |

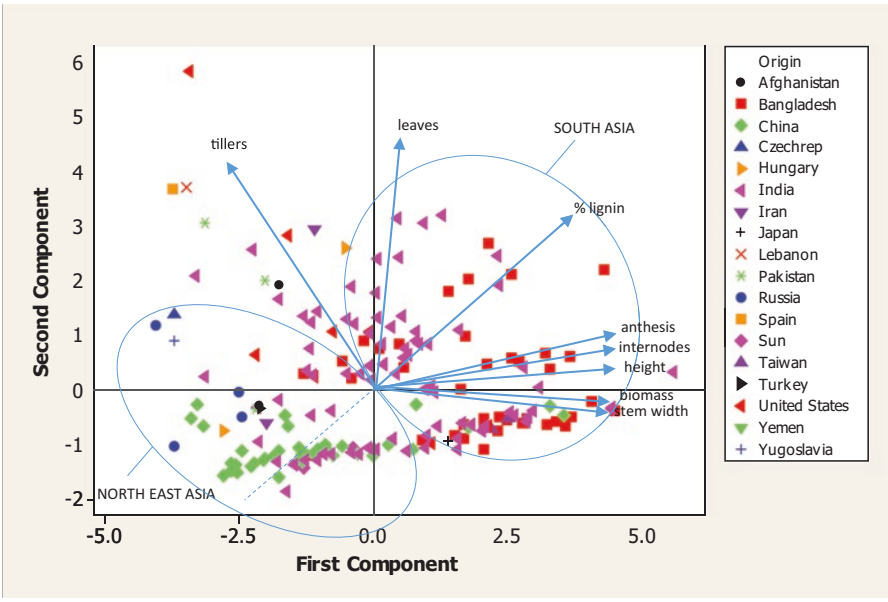


**Fig. 13.10** Geographical relationship to physiological attributes of *Setaria* genotypes. Principal component analysis (PCA) of physiological attributes, including lignin composition of stem tissue categorised by geographical origin of 214 *Setaria italica* germplasm lines collected from around the world. The trial was conducted at the University of Newcastle, NSW Australia under controlled temperature of 24/20 day/night and natural day-length (10.5–14.5 h)

attributes such as stem width and number of internodes were positively correlated to height and maturity (see PCA in Fig. 13.10; a narrow angle between vectors is indicative of high correlation). To a lesser extent percent lignin was also positively correlated to height and maturity whilst tiller number and leaf number traits in this germplasm set were independent of height and maturity.

With respect to percent lignin, the germplasm could be broadly divided into two groups, South Asia and North East Asia. Germplasm from South Asia (India, Bangladesh) was generally later maturing and had higher percent lignin compared to the group from North East Asia (China, Taiwan, Japan, Russia) that were earlier maturing and had lower percent lignin. Nevertheless, germplasm within the same maturity group varied greatly for percent lignin, and this material could be exploited to study lignin deposition.

### *13.7.1 Techniques for the Analysis of* Setaria *Cell Wall Traits*

To characterise the genetic factors that contribute to variation in cell wall traits, robust phenotyping systems are needed. Identification of traits that enhance the conversion potential of plant biomass into liquid fuels requires screening of thousands

of genetic variants. Phenotyping systems that are accurate, low cost and high-throughput are needed. Examples of phenotyping systems include enzymatic assays to evaluate saccharification efficiency of biomass, chromatographic tools to identify differences in composition of cell walls and spectroscopic methods such as near-infrared (NIR), Fourier-transform infrared (FTIR) and pyrolysis molecular beam mass spectroscopy (Torres et al. 2015). These techniques are combined with conventional chemical assays to analyse a set of samples and build calibration models to link compositional information with spectral variation (Martin et al. 2013). These models and screening tools are then used to efficiently predict the biochemical properties of unknown samples based on their spectral fingerprint.

### *13.7.2 Fourier Transform Infra Red Spectroscopy: Application of a Robust Tool for Cell Wall Characterisation*

Infrared (IR) spectroscopy can be readily applied as a rapid non-destructive tool for the investigation of cell wall composition in eudicots and monocot grasses. IR spectroscopy measures the transitions between molecular vibrational energy states as a result of absorption of mid-IR radiation. The vibrational energy levels are unique to each molecule and the frequencies of these molecular vibrations depend upon atomic mass, the geometric arrangement of the atoms and the strength of their chemical bonds (Larkin 2011). Mid-infrared (MIR) absorption spectra (4000–400 $cm^{-1}$; 25,000–2500 nm) provide characteristic fundamental vibrational modes directly, generating sharp peaks that are more readily interpreted than near-infrared (NIR; 13,000–4000 $cm^{-1}$; 700–2500 nm) or far-infrared (FIR; 10–400 $cm^{-1}$) spectra which reflect the broad overtones or combined bands of fundamental vibrations (Larkin 2011).

Although Fourier Transform MIR (FT-MIR or FTIR) spectra have been used in the analysis of plant cell walls for more than 30 years (Morikawa and Senda 1978), the power of this tool has only been realised more recently, when coupled with multivariate data analysis techniques. FTIR spectroscopy is a robust and accurate method for high-throughput screening of cell wall mutations in experimental plant tissues, such as the model species *Arabidopsis thaliana* (Mouille et al. 2003) and *Zea mays* coleoptiles (Carpita et al. 2001). In Sorghum, variation in the cell wall composition of a range of phenotypically different ecotypes grown in the field was identified following Principal Component Analysis (PCA) of FTIR spectra generated from ground stem samples (Fig. 13.11). The PCA separation of the different Sorghum ecotypes (Fig. 13.11b) is based predominantly upon the wavenumbers corresponding to cellulose and lignin as shown in Fig. 13.11c. Total lignin content was measured by the Acetyl Bromide method (Fig. 13.11d), validating the differences in cell wall composition, particularly lignin, predicted by FTIR and PCA.

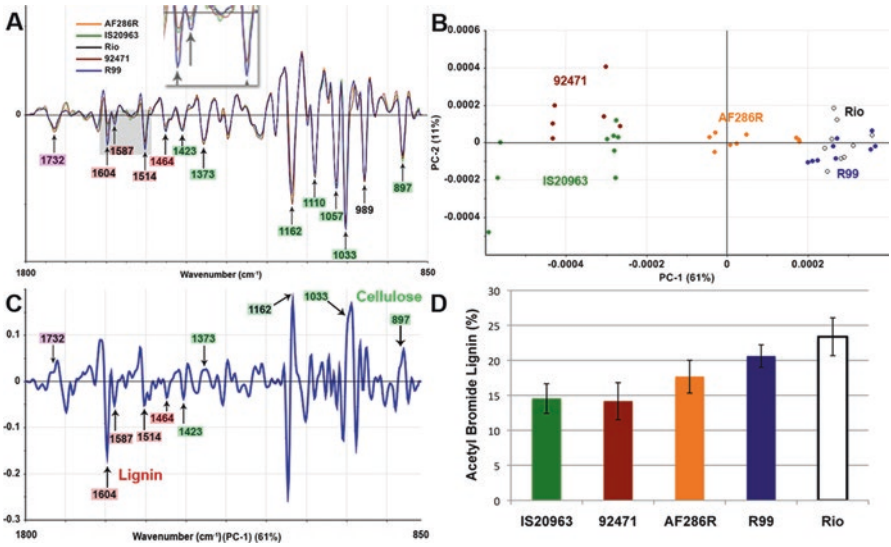


**Fig. 13.11** Principal component analysis (PCA) of FTIR spectra collected from *Sorghum bicolor* accessions grown under field conditions. (**a**) Second derivative spectra with extended multiplicative scatter correction (EMSC) applied; Inset highlights the differences between accessions at wavenumbers 1587 and 1604 $cm^{-1}$ (total lignin). (**b**) Scores plot of PC1 against PC2 (**c**) PC1 loadings plot showing the main peaks contributing to PC1 and (**d**) Total lignin as determined by the Acetyl Bromide method (Martin, Byrt, Furbank and Grof; unpublished) Accession names are IS20963, 92471, AF286R, R99 and Rio

### *13.7.3 Development of a Partial Least Squares (PLS) Model to Predict Cell Wall Composition in* Setaria

FTIR has also been used as a tool to assess *Setaria* as a model for unravelling the complexity of cell wall composition, assembly and subsequently deconstruction in grasses. Drawing upon a set of 214 *S. italica* accessions a prediction of total lignin was made with an accuracy of close to 90 % using FTIR spectra (Fig. 13.12). This research is currently being extended to incorporate the prediction of the full gamut of carbohydrate components making up *Setaria* cell walls following hydrolysis.

## 13.8 Conclusion

The gene regulatory network controlling cell wall synthesis and modification in the developing *Setaria* stem internode is not yet known. One strategy to identify this network is gene expression profiling. In addition to the genes involved in the biosynthesis of components of the cell wall, there are also upstream transcription

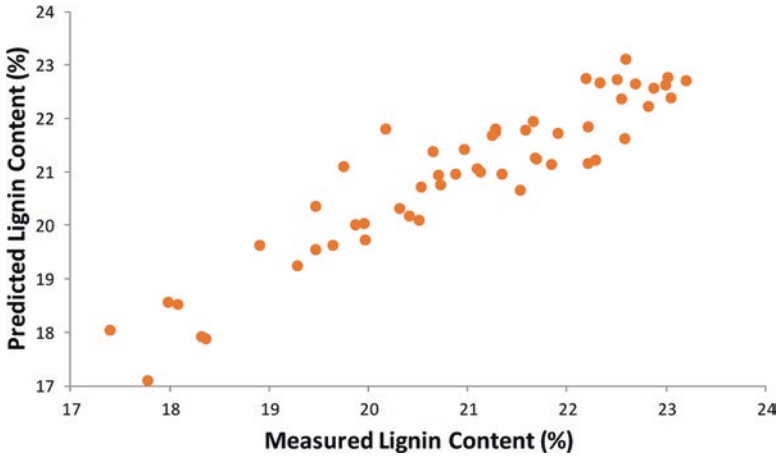


**Fig. 13.12** Prediction of the percentage of total lignin in accessions of *Setaria italica* using FTIR spectra. FTIR-based predictions of lignin content in *Setaria* stem tissue that had been ground to a fine powder were made using a 2 factor partial least squared (PLS) predictive model. These values were plotted against wet chemistry data for lignin performed on the same samples

factors influencing cell wall biosynthesis. Identification and verification of these 'master regulators' in grasses is lagging behind relative to information about model eudicot species. We have reported a high-throughput experimental platform that will aid in the discovery of key genes involved in photoassimilate partitioning in the largest contributor to monocot plant biomass, the culm. This experimental platform includes *Setaria* gene expression analysis and characterisation of differences in biomass composition of *Setaria* stems using FTIR. The genetic co-linearity of grasses (Devos 2005) ensures that orthologues of key candidate genes identified in *Setaria* can be rapidly identified in the closely related annual or perennial $C_4$ monocot species including economically important crop plants such as maize and Sorghum.

## References

Abreu HS, Latorraca JVF, Pereira RPW, Monteiro MBO, Abreu FA, Amparado KF. A supramolecular proposal of lignin structure and its relation with the wood properties. An Acad Bras Cienc. 2009;81:137–42.

Alejandro S, Lee Y, Tohge T, et al. AtABCG29 is a monolignol transporter involved in lignin biosynthesis. Curr Biol. 2012;22:1207–12.

Aohara T, Kotake T, Kaneko Y, Takatsuji H, Tsumuraya Y, Kawasaki S. Rice BRITTLE CULM 5 (BRITTLE NODE) is involved in secondary cell wall formation in the sclerenchyma tissue of nodes. Plant Cell Physiol. 2009;50:1886–97.

Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X, Kollipara K, Carrigan L, Tomes D, Chapman M, Dhugga K. Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. Cellulose. 2004;11:287–99.

Arioli T, Peng L, Betzner AS, et al. Molecular analysis of cellulose biosynthesis in Arabidopsis. Science. 1998;279:717–20.
Bar-Peled M, O'Neill MA. Plant nucleotide sugar formation, interconversion, and salvage by sugar recycling. Annu Rev Plant Biol. 2011;62:127–55.
Boerjan W, Ralph J, Baucher M. Lignin biosynthesis. Annu Rev Plant Biol. 2003;54:519–46.
Boija E, Johansson G. Interactions between model membranes and lignin-related compounds studied by immobilized liposome chromatography. Biochim Biophys Acta Biomembr. 1758; 2006:620–6.
Bootten TJ, Harris PJ, Melton LD, Newman RH. Solid-state 13C-NMR spectroscopy shows that the xyloglucans in the primary cell walls of mung bean (Vigna radiata L.) occur in different domains: a new model for xyloglucan-cellulose interactions in the cell wall. J Exp Bot. 2004;55:571–83.
Bosch M, Mayer CD, Cookson A, Donnison IS. Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. J Exp Bot. 2011;62:3545–61.
Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X-G, Kellogg E, Van Eck J. Setaria viridis: a model for C4 photosynthesis. Plant Cell. 2010;22:2537–44.
Bunzel M, Ralph J, Brüning P, Steinhart H. Structural identification of dehydrotriferulic and dehydrotetraferulic acids isolated from insoluble maize bran fiber. J Agric Food Chem. 2006;54:6409–18.
Burton RA, Wilson SM, Hrmova M, Harvey AJ, Shirley NJ, Medhurst A, Stone BA, Newbigin EJ, Bacic A, Fincher GB. Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-beta-D-glucans. Science. 2006;311:1940–2.
Campbell JA, Davies GJ, Bulone V, Henrissat B. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem J. 1997;326: 929–39.
Carpita NC. Hemicellulosic polymers of cell walls of Zea coleoptiles. Plant Physiol. 1983;72: 515–21.
Carpita NC. Update on mechanisms of plant cell wall biosynthesis: how plants make cellulose and other (1→4)-β-D-Glycans. Plant Physiol. 2011;155:171–84.
Carpita NC. Progress in the biological synthesis of the plant cell wall: new ideas for improving biomass for bioenergy. Curr Opin Biotechnol. 2012;23:330–7.
Carpita NC, Defernez M, Findlay K, Wells B, Shoue DA, Catchpole G, Wilson RH, McCann MC. Cell wall architecture of the elongating maize coleoptile. Plant Physiol. 2001;127:551–65.
Casu RE, Jarmey JM, Bonnett GD, Manners JM. Identification of transcripts associated with cell wall metabolism and development in the stem of sugarcane by Affymetrix GeneChip Sugarcane Genome Array expression profiling. Funct Integr Genomics. 2007;7:153–67.
Cheynier V, Comte G, Davies KM, Lattanzio V, Martens S. Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. Plant Physiol Biochem. 2013;72:1–20.
Collins HM, Burton RA, Topping DL, Liao M-L, Bacic A, Fincher GB. Variability in fine structures of noncellulosic cell wall polysaccharides from cereal grains: potential importance in human health and nutrition. Cereal Chem. 2010;87:272–82.
Cosgrove DJ. Growth of the plant cell wall. Nat Rev Mol Cell Biol. 2005;6:850–61.
Dai J, Mumper RJ. Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. Molecules. 2010;15:7313–52.
de Oliveira Buanafina MM. Feruloylation in grasses: current and future perspectives. Mol Plant. 2009;2:861–72.
de Oliveira Buanafina MM, Cosgrove DJ. Cell walls: structure and biogenesis. In: Sugarcane: physiology, biochemistry and functional biology. New York: Wiley; 2013. p. 307–329.
Devos KM. Updating the "crop circle". Curr Opin Plant Biol. 2005;8:155–62.
Doblin MS, Pettolino FA, Wilson SM, Campbell R, Burton RA, Fincher GB, Newbigin E, Bacic A. A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-β-D-glucan synthesis in transgenic Arabidopsis. Proc Natl Acad Sci U S A. 2009;106:5996–6001.

Doblin MS, Pettolino F, Bacic A. Evans review : plant cell walls: the skeleton of the plant world. Funct Plant Biol. 2010;37:357–81.
Doering A, Lathe R, Persson S. An update on xylan synthesis. Mol Plant. 2012;5:769–71.
Ehlting J, Mattheus N, Aeschliman DS, et al. Global transcript profiling of primary stems from Arabidopsis thaliana identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. Plant J. 2005;42:618–40.
Emons AMC, Mulder BM. The making of the architecture of the plant cell wall: how cells exploit geometry. Proc Natl Acad Sci U S A. 1998;95:7215–9.
Faik A. Xylan biosynthesis: news from the grass. Plant Physiol. 2010;153:396–402.
Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperley DC, Kennedy CJ, Jarvis MC. Nanostructure of cellulose microfibrils in spruce wood. Proc Natl Acad Sci U S A. 2011;108:1195–203.
Hansen SF, Harholt J, Oikawa A, Scheller HV. Plant glycosyltransferases beyond CAZy: a perspective on DUF families. Front Plant Sci. 2012;3:59.
Hao Z, Mohnen D. A review of xylan and lignin biosynthesis: foundation for studying Arabidopsis irregular xylem mutants with pleiotropic phenotypes. Crit Rev Biochem Mol Biol. 2014;49:212–41.
Harholt J, Suttangkakul A, Scheller HV. Biosynthesis of pectin. Plant Physiol. 2010;153:384–95.
Hatfield R, Vermerris W. Lignin formation in plants. The dilemma of linkage specificity. Plant Physiol. 2001;126:1351–7.
He L, Terashima N. Formation and structure of lignin in monocotyledons IV. Deposition process and structural diversity of the lignin in the cell wall of sugarcane and rice plant studied by ultraviolet microscopic spectroscopy. Holzforsch Int J Biol Chem Phys Technol Wood. 1991;45:191.
Hill JL, Hammudi MB, Tien M. The Arabidopsis cellulose synthase complex: a proposed hexamer of CESA trimers in an equimolar stoichiometry. Plant Cell. 2014;26:4834–42.
Hirano K, Aya K, Morinaka Y, Nagamatsu S, Sato Y, Antonio BA, Namiki N, Nagamura Y, Matsuoka M. Survey of genes involved in rice secondary cell wall formation through a co-expression network. Plant Cell Physiol. 2013;54:1803–21.
Jacquet G, Pollet B, Lapierre C, Mhamdi F, Rolando C. New ether-linked ferulic acid-coniferyl alcohol dimers identified in grass straws. J Agric Food Chem. 1995;43:2746–51.
Kaneda M, Rensing KH, Wong JCT, Banno B, Mansfield SD, Samuels AL. Tracking monolignols during wood development in Lodgepole pine. Plant Physiol. 2008;147:1750–60.
Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, Martinoia E. Plant ABC transporters. Arabidopsis Book. 2011;9:e0153.
Kim J-B, Olek AT, Carpita NC. Cell wall and membrane-associated Exo-β-d-glucanases from developing maize seedlings. Plant Physiol. 2000;123:471–86.
Kimura S, Laosinchai W, Itoh T, Cui X, Linder CR, Brown RM. Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant vigna angularis. Plant Cell. 1999;11:2075–86.
Kokubo A, Sakurai N, Kuraishi S, Takeda K. Culm brittleness of barley (Hordeum vulgare L.) mutants is caused by smaller number of cellulose molecules in cell wall. Plant Physiol. 1991;97:509–14.
Koutaniemi S. Lignin biosynthesis in Norway spruce: from a model system to the tree. 2007. Ph.D. thesis, University of Helsinki.
Larkin P. Infrared and Raman spectroscopy; principles and spectral interpretation. 1st ed.; 2011. ISBN 9780123869845.
Lattanzio V, Lattanzio VMT, Cardinali A. Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects. In: Phytochemistry: advances in research. Kerala, India: Research Signpost; 2006. p. 23–45.
Li X, Weng JK, Chapple C. Improvement of biomass through lignin modification. Plant J. 2008;54:569–81.

Liu CJ. Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly. Mol Plant. 2012;5:304–17.
Martin AP, Palmer WM, Byrt CS, Furbank RT, Grof CP. A holistic high-throughput screening framework for biofuel feedstock assessment that characterises variations in soluble sugars and cell wall composition in Sorghum bicolor. Biotechnol Biofuels. 2013;6:186.
Martin AP, Palmer WM, Brown C, Abel C, Lunn JE, Furbank RT, Grof CPL. A developing Setaria viridis internode: an experimental system for the study of biomass generation in a C4 model species. Biotechnol Biofuels. 2016;9:45.
McCann MC, Carpita NC. Designing the deconstruction of plant cell walls. Curr Opin Plant Biol. 2008;11:314–20.
McFarlane HE, Döring A, Persson S. The cell biology of cellulose synthesis. Annu Rev Plant Biol. 2014;65:69–94.
Mitchell RA, Dupree P, Shewry PR. A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. Plant Physiol. 2007;144:43–53.
Morikawa H, Senda M. Infrared analysis of oat coleoptile cell walls and oriented structure of matrix polysaccharides in the walls. Plant Cell Physiol. 1978;19:327–36.
Mortimer JC, Miles GP, Brown DM, et al. Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. Proc Natl Acad Sci U S A. 2010;107:17409–14.
Mouille G, Robin S, Lecomte M, Pagant S, Hofte H. Classification and identification of Arabidopsis cell wall mutants using Fourier-Transform InfraRed (FT-IR) microspectroscopy. Plant J. 2003; 35:393–404.
Mutwil M, Debolt S, Persson S. Cellulose synthesis: a complex complex. Curr Opin Plant Biol. 2008;11:252–7.
Nakashima J, Chen F, Jackson L, Shadle G, Dixon RA. Multi-site genetic modification of monolignol biosynthesis in alfalfa (Medicago sativa): effects on lignin composition in specific cell types. New Phytol. 2008;179:738–50.
Niimura H, Yokoyama T, Kimura S, Matsumoto Y, Kuga S. AFM observation of ultrathin microfibrils in fruit tissues. Cellulose. 2010;17:13–8.
Oehme DP, Downton MT, Doblin MS, Wagner J, Gidley MJ, Bacic A. Unique aspects of the structure and dynamics of elementary Iβ cellulose microfibrils revealed by computational simulations. Plant Physiol. 2015;168:3–17.
Parameswaran N, Liese W. On the fine structure of bamboo fibres. Wood Sci Technol. 1976; 10:231–46.
Park YB, Cosgrove DJ. A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. Plant Physiol. 2012;158:1933–43.
Park YB, Cosgrove DJ. Xyloglucan and its interactions with other components of the growing cell wall. Plant Cell Physiol. 2015;56:180–94.
Paterson AH, Bowers JE, Bruggmann R, et al. The Sorghum bicolor genome and the diversification of grasses. Nature. 2009;457:551–6.
Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G. Hemicellulose biosynthesis. Planta. 2013;238:627–42.
Penning BW, Hunter CT, Tayengwa R, et al. Genetic resources for maize cell wall biology. Plant Physiol. 2009;151:1703–28.
Ralph J, Grabber JH, Hatfield RD. Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. Carbohydr Res. 1995;275:167–78.
Rennie EA, Scheller HV. Xylan biosynthesis. Curr Opin Biotechnol. 2014;26:100–7.
Richmond T. Higher plant cellulose synthases. Genome Biol. 2000;1:reviews 3001.1–3001.6.
Richmond TA, Somerville CR. The cellulose synthase superfamily. Plant Physiol. 2000;124: 495–8.
Saha J, Sengupta A, Gupta K, Gupta B. Molecular phylogenetic study and expression analysis of ATP-binding cassette transporter gene family in Oryza sativa in response to salt stress. Comput Biol Chem. 2015;54:18–32.

Sattler SE, Funnell-Harris DL. Modifying lignin to improve bioenergy feedstocks: strengthening the barrier against pathogens? Front Plant Sci. 2013;4:70. doi:10.3389/fpls.2013.00070.
Scheller HV, Ulvskov P. Hemicelluloses. Annu Rev Plant Biol. 2010;61:263–89.
Sibout R, Höfte H. Plant cell biology: the ABC of monolignol transport. Curr Biol. 2012;22: 533–5.
Sindhu A, Langewisch T, Olek A, Multani DS, McCann MC, Vermerris W, Carpita NC, Johal G. Maize Brittle stalk2 encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity. Plant Physiol. 2007;145:1444–59.
Tanaka K, Murata K, Yamazaki M, Onosato K, Miyao A, Hirochika H. Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. Plant Physiol. 2003;133:73–83.
Taylor NG. Cellulose biosynthesis and deposition in higher plants. New Phytol. 2008;178: 239–52.
Torres AF, Visser RGF, Trindade LM. Bioethanol from maize cell walls: genes, molecular tools, and breeding prospects. GCB Bioenerg. 2015;7:591–607.
Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. Lignin biosynthesis and structure. Plant Physiol. 2010;153:895–905.
Vogel J. Unique aspects of the grass cell wall. Curr Opin Plant Biol. 2008;11:301–7.
Wilkerson CG, Mansfield SD, Lu F, et al. Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. Science. 2014;344:90–3.
Wilson SM, Ho YY, Lampugnani ER, Van de Meene AML, Bain MP, Bacic A, Doblin MS. Determining the subcellular location of synthesis and assembly of the cell wall polysaccharide (1,3; 1,4)-β-D-Glucan in grasses. Plant Cell. 2015;27:754–71.
Xu J, Li Y, Ma X, Ding J, Wang K, Wang S, Tian Y, Zhang H, Zhu XG. Whole transcriptome analysis using next-generation sequencing of model species Setaria viridis to support C4 photosynthesis research. Plant Mol Biol. 2013;83:77–87.
York WS, O'Neill MA. Biochemical control of xylan biosynthesis—which end is up? Curr Opin Plant Biol. 2008;11:258–65.
Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A. A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. Plant Physiol. 2010;154:78–97.
Zhang QS, Cheetamun R, Dhugga KS, et al. Spatial gradients in cell wall composition and transcriptional profiles along elongating maize internodes. BMC Plant Biol. 2014;14:27–46. doi:10.1186/1471-2229-14-27.
Zhong R, Ye Z-H. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. Plant Cell Physiol. 2015;56:195–214.

# Chapter 14
# *Setaria* Root–Microbe Interactions

**Fernanda Plucani do Amaral, Beverly Jose Agtuca, and Gary Stacey**

**Abstract** Plants interact with a wide range of soil microorganisms. In some cases, this interaction can result in significant benefits to both microbe and plant host. For example, it is well established that some soil bacteria promote the growth of plants, which can result in significant increases in crop yield. The effects of such "plant growth-promoting rhizobacteria" (PGPR) has been demonstrated with a wide variety of plant species. However, the molecular mechanism behind this growth promotion is still largely undefined. As demonstrated in other experimental systems, we believe that adoption of a model plant species for study of PGPR–plant interactions would promote detailed mechanistic studies and ultimately lead to broader and more effective use of such bacteria in agricultural production.

**Keywords** Rhizosphere microbiota • Plant growth-promoting bacteria • *Setaria viridis* • Grass model system • Root bacteria interaction

## 14.1 Introduction

The rhizosphere, the narrow region of soil that surrounds and is influenced by the plant root, is home to a wide range of microbes. These microbes interact naturally between themselves, as well as with the plant, forming a close and complex communication network. For instance, this communication network can influence the host plant's growth, nutrition, and health (Bonfante and Anca 2009; Mendes et al. 2011; Berendsen et al. 2012). The rhizosphere microbiota can also increase the biomass and composition of plant communities, as well as plant interactions with antagonistic and mutualistic symbionts in natural ecosystems (Kardol et al. 2007; Schnitzer et al. 2011). However, the molecular mechanisms defining how these microbes interact with plants are still unclear. Some insights have come from studies of model $C_3$ plant species, such as rice (de Souza et al. 2013; de Brito Ferreira et al.

F.P. do Amaral, Ph.D. • B.J. Agtuca, B.S. • G. Stacey, Ph.D. (✉)
Divisions of Plant Science and Biochemistry, University of Missouri,
201 Christopher S. Bond Life Sciences Center, 1201 E. Rollins Street,
Columbia, MO 65211, USA
e-mail: doamaralf@missouri.edu; bjamtf@mail.missouri.edu; staceyg@misouri.edu



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_14

2014) and *Arabidopsis* (Bulgarelli et al. 2012; Lundberg et al. 2012; Bodenhausen et al. 2013), as well as a few $C_4$ plant species including maize (Amaral et al. 2014; Li et al. 2014) sorghum (Gopalakrishnan et al. 2013), switchgrass (Mao et al. 2014), and sugarcane (Oliveira et al. 2009; Vargas et al. 2014).

With regard to $C_4$ plant species, the suggestion has been made that *Setaria viridis* is a potentially excellent model due to its short life cycle, small stature, large seed production, small genome size, and potential for genetic transformation (Li and Brutnell 2011; Kumar et al. 2013; Jiang et al. 2013). In addition to supporting genetic and molecular studies, *S. viridis* also has the potential to serve as a useful model to explore the molecular details of rhizosphere microbe colonization, infection, growth and maintenance, and plant growth promotion and maintenance. There is great interest, for example, in using plant growth-promoting rhizobacteria for sustainable agricultural production, especially for bioenergy crops. Such uses could contribute to mitigation of climate change effects. In order to fully implement such strategies, much more needs to be known about the ecology, composition, dynamics, and activities of rhizosphere microbial communities; information that could be more rapidly and efficiently obtained through widespread adoption of a plant model. There are strong arguments that *S. viridis* can and should fulfill this role.

## 14.2 Plant–Rhizosphere Microbiota Interactions

Root growth and activity can change the physical, chemical, and biological properties of the rhizosphere. It is well known that root exudation plays an important role in the establishment of plant–microorganism interactions. These exudates provide a wide variety of organic compounds that can act as signaling molecules, nutrients, inhibitors (e.g., phytoalexins), phytohormones, enzymes, or allelochemicals (Nannipieri et al. 2007). These compounds allow the root system to impact the rhizosphere microbiota, allowing them to interact and display fundamental roles of recognition and colonization (Grayston et al. 1997). Plant exudates can serve as a carbon source for the microbes, as well as attracting phytobeneficial soil microbes, whose presence influences plant development (e.g., rooting patterns), nutrient availability, and pathogen persistence in the rhizosphere (Bolton et al. 1993; Bowen and Rovira 1999; Barea 2000; Haichar et al. 2008, 2014). Root exudation also provides physical benefits to the plant such as a reduction in root desiccation, reduction of friction between root tips and soil, and improvement of the structural stability of soil (Rougier and Chaboud 1985). Thus, it is now clearly established that root exudation is an important mediator in belowground plant–microorganism interactions in the rhizosphere. However, it is important to consider that the rate, composition, and extent of exudations depend on genetic factors and differ extensively among plant species and environmental conditions (Kochian et al. 2005).

There are many biotic and abiotic factors that can affect microbial community assembly in the rhizosphere. Many studies have shown that "soil properties," as an example of an abiotic factor, influences the assembly of bacterial communities in the rhizosphere,

plant physiology, and root exudation patterns (de Ridder-Duine et al. 2005; Andrew et al. 2012; Inceoglu et al. 2012). This is due to the soil's complex physicochemical characteristics such as pH, salinity, texture, organic matter content, concentration of nutrient elements, and seasonal effects, as well as management practices like irrigation, tillage, cropping, fertilizer and pesticide application, and residue incorporation (Grayston et al. 1997; Macdonald et al. 2004; Fang et al. 2005; Ibekwe et al. 2010). Studies using sequence-based analyses of the rhizosphere microbiome of different *A. thaliana* ecotypes demonstrated that the soil type had a significant influence on microbiota composition (Bulgarelli et al. 2012; Lundberg et al. 2012). The microbial community in the rhizosphere of maize plants is also influenced by soil type (Chiarini et al. 1998). The plant genotype can significantly influence the composition and activity of rhizosphere bacteria. It can also affect root morphology, as well as the amount and type of root exudation (Bergsma-Vlami et al. 2005; Pivato et al. 2007; Bressan et al. 2009; Ladygina and Hedlund 2010). Such differences have been found in comparisons between plant species, as well as different genotypes of the same species (Inceoglu et al. 2012; Hardoim et al. 2011; Mazzola et al. 2004; Yao and Wu 2010). An example is the characterization of the rhizosphere microbiota of three wheat cultivars grown at two distant field sites. The results showed that microbial community composition was dependent on the cultivar (Siciliano et al. 1998). In contrast, different maize cultivars did not show any significant difference in bacterial community structures (Chiarini et al. 1998; Dalmastri et al. 1999). Plant age and developmental stage can also affect the diversity of the rhizosphere microbial community (Inceoglu et al. 2012; Brimecombe et al. 2001; Dumbrell et al. 2011). A young plant contains r-strategy organisms in the rhizosphere, bacterial species that have fast growth rates, while an older plant has k-strategy organisms, bacterial species that have slower growth rates in the rhizosphere (Brimecombe et al. 2001; Folman et al. 2001). Overall, it may be said that plant–rhizosphere microbiota interactions and diversity are determined by many interrelated biotic and abiotic factors.

Such genotype-specific effects on rhizosphere microbial populations have also been noted in studies using *S. viridis*. Li et al. (2014) studied root biomass and colonization of *S. viridis* inoculated with deleterious rhizobacteria, showing that soils with differing pH and amounts of organic matter impacted *Setaria* growth relative to uninoculated controls. More recently, Pankievicz et al. (2015) screened over 30 genotypes of *S. viridis* for their response to inoculation with two Plant Growth-Promoting Rhizobacteria (PGPR) in comparison to uninoculated controls. This screen identified only three genotypes that significantly and consistently responded to inoculation. Therefore, as in other studies, the *S. viridis* rhizosphere microbiota is influenced both by plant genotype and soil type.

## 14.3 Beneficial Bacterial Interactions

Beneficial interactions between plant and microorganisms include intimate, symbiotic associations (e.g., rhizobia and mycorrhizae), as well as less intimate interactions (e.g., with PGPR) (Walker et al. 2012). These interactions can be harmful, neutral,

or beneficial according to how they impact the plant. Among the beneficial plant bacterial associations, PGPR stand out due to their ability to enhance plant growth (Martinez-Viveros et al. 2010; Prashar et al. 2014). PGPR can act as (1) bioprotectants, (2) biostimulants, and/or (3) biofertilizers (Murphy et al. 2000; Zamioudis and Pieterse 2012). As bioprotectants, PGPR can control plant diseases (Zehnder et al. 1997) and induce a systemic resistance response in the plant that is naturally effective against multiple pathogens (Cameron et al. 1994), insect pests (Wei et al. 1996), fungi (Hu et al. 2009), bacteria (Burkett-Cadena et al. 2008; Spaepen et al. 2008), and nematodes (Spaepen and Vanderleyden 2011). As biostimulants, PGPR are able to produce phytohormones such as auxin, abscisic acid, and cytokinin (Steenhoudt and Vanderleyden 2000; Okon and Labanderagonzalez 1994). These phytohormones can significantly impact the plant root system, for example, stimulating root elongation and branching (Dobbelaere et al. 2001).

As biofertilizers, PGPR enhance the plant's ability to uptake nutrients, and thereby increase crop yields (Glick 1995; Berg 2009). There are many examples of PGPR directly increasing crop yields including *Enterobacter sp.* in maize, *Gluconacetobacter diazotrophicus* in sugarcane, *Burkholderia sp.* in rice and maize, *Herbaspirillum seropedicae* in rice, sugarcane, and wheat, and *Azospirillum brasilense* in maize, sorghum, wheat, and pearl millet (Carvalho et al. 2014). What is less clear is the mechanism by which PGPR increase crop yields. For example, in some cases these effects have been attributed to biological nitrogen fixation (BNF) (Mitter et al. 2013) or phosphate solubilization (Rodriguez and Fraga 1999). A BNF ability seems to be rather common among PGPR, and it is presumed that this fixed nitrogen is ultimately available to increase plant growth (Mitter et al. 2013). Given that nitrogen is often a limiting nutrient in soil, BNF ability can be an important source of nitrogen input in agricultural systems; such as is the case for the legume-rhizobium symbiosis. Also common among PGPR is the ability to solubilize phosphate. Examples include *Azospirillum*, *Burkholderia*, and *Enterobacter* that can convert insoluble phosphorous into a soluble form through acidification due to secretion of organic acids. These bacteria can significantly improve the phosphate nutrition of the colonized host plant (Santi et al. 2013; da Costa et al. 2013).

## 14.4 The Case for a Plant Model to Study the Mechanisms of PGPR Action

As discussed above, there are a variety of possible ways by which PGPR might promote plant growth but there are few examples in which the actual mechanism has been elucidated beyond the level of initial observation. Biological nitrogen fixation is an excellent example in which the ability of PGPR to fix nitrogen has been well established, as well as the ability to conduct BNF *in planta*. However, less clear is the actual ability of the plant to benefit from this fixed nitrogen. This and other examples point to the need for a genetically tractable model system in which the mechanism of action of PGPR can be studied in detail. Arabidopsis is clearly one

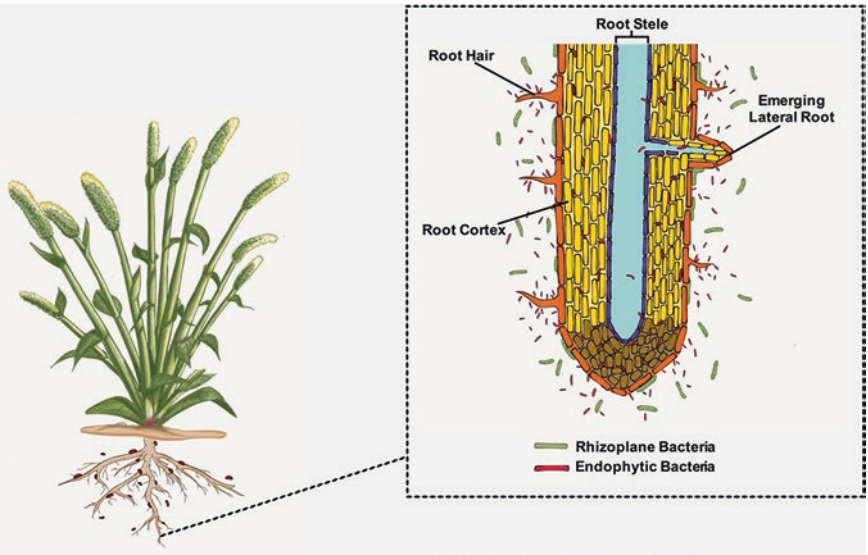


**Fig. 14.1** Colonization on *Setaria viridis* roots by diazotrophic PGPR. Rhizoplane bacteria (*green cells*) colonize the soil area and the root surface without invading the root's internal tissues. In contrast, endophytic bacteria (*red cells*) have the ability to colonize the intercellular spaces of the host's root. Endophytic bacteria penetrate their hosts through the discontinuities of the epidermis, such as sites of lateral root emergence

choice for such a model. However, given the expected application of PGPR in grass cropping systems (e.g., bioenergy crops), it seems wise to also consider a $C_4$ grass species. It is for this reason that Pankievicz et al. (2015) recently adopted *S. viridis* as a model to study the interaction with two well-characterized PGPR species, *Herbaspirillum seropedicae* and *Azospirillum brasilense*. Plants inoculated with *H. seropedicae* RAM4, expressing the DsRed fluorescent protein (Monteiro et al. 2008), showed that this bacteria colonizes the intercellular spaces of *S. viridis* roots, forming micro colonies, preferentially at sites of lateral root emergence. Indeed, expression of β-glucuronidase from a *nifH-gusA* fusion in *A. brasilense* strain FP2-7 (Machado et al. 1991) showed *S. viridis* root colonization by this bacterium and also nitrogenase gene expression, primarily on the surface of roots tips and the elongation zone. *H. seropedicae*, isolated from rice, is a well-characterized endophytic and diazotrophic bacteria that can internally colonize the roots of a variety of plant species without causing visible, harmful effects (Baldani et al. 1992; Pedrosa et al. 2011). In contrast, *A. brasilense*, isolated from wheat and maize, has been shown to primarily colonize the exterior of plant roots (Tarrand et al. 1978; Bashan et al. 2004; Xu et al. 2013). Figure 14.1 illustrates the pattern of bacterial colonization of both the exterior and internal spaces of *S. viridis* roots. Besides colonization, Pankievicz et al. (2015) demonstrated that inoculation with both *H. seropedicae* and *A. brasilense* promoted plant growth as measured by a variety of parameters (e.g., root length and total biomass).

## 14.5 *Setaria* Root Bacteria Interaction

$C_4$ plants, such as maize and sorghum, usually have a higher water and nitrogen use efficiency and photosynthetic efficiency compared to $C_3$ plants (Xu et al. 2013). These factors can increase overall crop productivity of many important food crops and bioenergy grasses (Brutnell et al. 2010; Sage et al. 2012).

*Setaria viridis* (green foxtail), the weedy relative of *S. italica* (foxtail millet), also possesses attributes suitable for genetic analyses including a small stature, rapid life cycle, prolific seed production, and a relatively small and sequenced genome (Bennetzen 2012). *Setaria* species are also morphologically similar to most Panicoideae grasses, including potential biofuel feedstocks, such as switchgrass (*Panicum virgatum*) and Miscanthus (*Miscanthus giganteus*) (Brutnell et al. 2012). Therefore, understanding the mechanisms underpinning growth promotion by bacteria associations in this $C_4$ model grass system could have far reaching benefits to both food and bioenergy production.

The relationship between *Setaria* and bacteria of the genus *Azospirillum* has previously been exploited in studies of nitrogen incorporation. Raffi et al. (2012) isolated 20 *Azospirillum* spp. from the roots and soil of foxtail millet. All of these culture isolates were found to be diazotrophic suggesting the potential for high levels of nitrogen input from BNF. Among the 20 isolates, the most common species was *Azospirillum lipoferum*. Indeed, inoculation of *Setaria* with *A. brasilense* was found to significantly increase the dry weight of roots and shoots under nitrogen limiting growth conditions Okon et al. (1983). The authors used $^{15}N_2$ to demonstrate that approximately 5% of the nitrogen fixed by *A. brasilense* was incorporated into *S. italica* roots. In agreement, Pankievicz et al. (2015) used $^{13}$NN radioisotope labeling to estimate that approximately 7 % of the nitrogen fixed by *A. brasilense* and *H. seropedicae,* was incorporated by the plant. Indeed, the $^{13}$N label could be directly traced to plant protein (i.e., ribulose 1,5 bisphosphate carboxylase). However, inoculation of *S. viridis* with *A. brasilense* strain HM053 (Machado et al. 1991), which fixes higher levels of nitrogen and excretes ammonia, resulted in fixation at levels that could provide 100 % of the daily nitrogen demands of the plant. This was reflected by the very robust growth of plants inoculated with strain HM053 under nitrogen limiting conditions. This study clearly demonstrated the utility of *S. viridis* to study BNF by PGPR, setting the stage for more mechanistic studies.

Bacterial root colonization involves migration towards the plant roots, absorption and anchoring onto the root surface, as well as subsequent bacterial proliferation (Reinhold et al. 1986; Compant et al. 2010). These steps might be stimulated by the direct promotion of root growth and branching by PGPR. For example, roots of *S. viridis* inoculated with PGPR, growing under severe nitrogen limitation, showed a 27 % increase in root length and 39 % increase in lateral root number relative to the uninoculated control plants Pankievicz et al. (2015). Indeed, plants grown under nitrogen starvation, when inoculated with *A. brasilense* strain HM053, showed an increase in growth parameters either without nitrate or with low nitrate addition (0.5 mM $KNO_3$), as shown in the Fig. 14.2a and b.

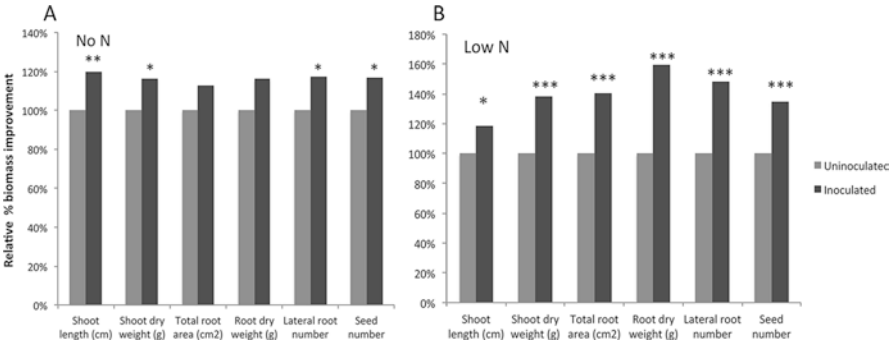


**Fig. 14.2** Relative percentage of biomass improvement of *S. viridis* A10.1 plants, inoculated with A. brasilense strain HM053 grown under two different nitrate condition. (**a**) Inoculated plants grown without nitrate addition (No N) showed significant growth promotion in shoot, lateral root number, and seed production 30 days after inoculation. (**b**) Inoculated plants grown with low nitrate addition (0.5 mM $KNO_3$) showed an increase in all growth parameters analyzed 30 days after inoculation. Asterisks represent a statistically significant difference **P* value≤0.05; ***P* value≤0.01; ****P* value≤0.001

These results demonstrate the resistance and plasticity of *Setaria* root responses under stress conditions. The association with PGPR influenced the host's ability to sustain healthy growth under nitrogen limitation, which clearly affected plant metabolism. For instance, $CO_2$ fixation was increased in inoculated plants relative to controls (Pankievicz et al. 2015). In addition, PGPR inoculation directly impacted the relative levels of various metabolite pools such that inoculated plants under nitrogen limitation were more similar to plants grown under nitrogen sufficiency than to the uninoculated, nitrogen-deprived controls. Overall, these data suggest that beneficial bacterial associations play an important role in enhancing plant development not only in roots, but through whole plant effects on growth and metabolism.

Since *Setaria* is a weed, it could also be used to examine the ability of bacteria to act as a bioherbicide. For example, *Pseudomonas fluorescens* BRG100, isolated from the rhizosphere of *S. viridis,* demonstrated herbicidal activity through the production of secondary metabolites when tested on other gramineaceous weeds. The metabolites were found to suppress weed root growth, mainly reducing root length (Caldwell et al. 2012).

In general, *Setaria* roots demonstrate high plasticity and an easy adaptation to stressful environments, and maintain plant growth through different strategies. It seems that the root system uses different adaptive mechanisms that change architecture and physiological characteristics in order to survive. Even though *Setaria* is a small plant, it has the potential to become a robust model system to explore grass–microbe interaction mechanisms.

## 14.6 Conclusions

Beneficial plant-associated bacteria play a key role in supporting and increasing plant health and growth through mechanisms that remain largely undefined. Successful plant growth promotion is dependent on the genotype of the plant host and PGPR, as well as a variety of physicochemical parameters. Hence, it is often difficult to predict the success of PGPR inoculations from one field to the next. We believe that greater consistency in inoculant performance and plant growth promotion will come when more is known about the mechanisms by which these interesting bacteria colonize and affect plant growth and development. Past examples of the application of model organisms to studies of complex systems (e.g., yeast, Drosophila, Arabidopsis, and mouse) suggest that adoption of a plant model suitable for the study of PGPR would provide significant benefits. Given the importance of grass species, especially $C_4$ grasses, to food and fuel production, *S. viridis* provides an attractive, genetically tractable model plant system. The available data suggest that this plant is readily colonized both in the lab and field by PGPR and, under the appropriate conditions, can demonstrate a robust response to inoculation. These initial studies lay the groundwork for more detailed mechanistic studies that can eventually provide the understanding to allow greater use of PGPR in cropping systems with the ultimate goal of more sustainable and profitable agriculture.

**Acknowledgements** The authors would like to acknowledge Johann Oliver Zinamon for drawing the Setaria-rhizosphere figure. Financial support for the research discussed from the authors' laboratory was provided by grant DE—FOA—0000223 (to G. Stacey) from the Department of Energy, Office of Biological and Environmental Research.

## References

Amaral FP, Bueno JCF, Hermes VS, Arisi ACM. Gene expression analysis of maize seedlings (DKB240 variety) inoculated with plant growth promoting bacterium *Herbaspirillum seropedicae*. Symbiosis. 2014;62:41–50.

Andrew DR, Fitak RR, Munguia-Vega A, Racolta A, Martinson VG, Dontsova K. Abiotic factors shape microbial diversity in Sonoran desert soils. Appl Environ Microbiol. 2012;78(21): 7527–37.

Baldani VLD, Baldani JI, Olivares F, Dobereiner J. Identification and ecology of *Herbaspirillum-seropedicae* and the closely related *Pseudomonas-rubrisubalbicans*. Symbiosis. 1992;13(1–3): 65–73.

Barea JM. Rhizosphere and mycorrhiza of field crops. In: Biological resource management: connecting science and policy. Berlin: Springer; 2000. p. 81–92.

Bashan Y, Holguin G. Proposal for the division of plant growth-promoting rhizobacteria into two classifications: biocontrol-PGPR (plant growth-promoting bacteria) and PGPR. Soil Biol Biochem. 1998;30(8–9):1225–8.

Bashan Y, Holguin G, de-Bashan LE. Azospirillum-plant relationships: physiological, molecular, agricultural, and environmental advances (1997–2003). Can J Microbiol. 2004;50(8):521–77.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30(6):555–61.
Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. Trends Plant Sci. 2012;17(8):478–86.
Berg G. Plant–microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. Appl Microbiol Biotechnol. 2009;84(1):11–8.
Bergsma-Vlami M, Prins ME, Raaijmakers JM. Influence of plant species on population dynamics, genotypic diversity and antibiotic production in the rhizosphere by indigenous *Pseudomonas* spp. Fems Microbiol Ecol. 2005;52(1):59–69.
Bodenhausen N, Horton MW, Bergelson J. Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. PLoS One. 2013;8(2):e56329.
Bolton H, Jr, Fredrickson JK, Elliott LF. Microbial ecology of the rhizosphere. In: Soil microbial ecology: applications in agricultural and environmental management. New York: Marcel Dekker; 1993. p. 27–63.
Bonfante P, Anca I-A. Plants, mycorrhizal fungi, and bacteria: a network of interactions. Annu Rev Microbiol. 2009;63:363–83.
Bowen GD, Rovira AD. The rhizosphere and its management to improve plant growth. Adv Agron. 1999;66:1–102.
Bressan M, Roncato M-A, Bellvert F, Comte G, Haichar FZ, Achouak W, et al. Exogenous glucosinolate produced by *Arabidopsis thaliana* has an impact on microbes in the rhizosphere and plant roots. ISME J. 2009;3(11):1243–57.
Brimecombe MJ, De Leij FA, Lynch JM. The effect of root exudates on rhizosphere microbial populations. In: Rhizosphere. New York: Marcel Dekker; 2001. p. 95–140.
Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X-G, et al. *Setaria viridis*: a model for C-4 photosynthesis. Plant Cell. 2010;22(8):2537–44.
Bulgarelli D, Rott M, Schlaeppi K, van Themaat EVL, Ahmadinejad N, Assenza F, et al. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature. 2012;488(7409):91–5.
Burkett-Cadena M, Kokalis-Burelle N, Lawrence KS, van Santen E, Kloepper JW. Suppressiveness of root-knot nematodes mediated by rhizobacteria. Biol Contr. 2008;47(1):55–9.
Caldwell CJ, Hynes RK, Boyetchko SM, Korber DR. Colonization and bioherbicidal activity on green foxtail by *Pseudomonas fluorescens* BRG100 in a pesta formulation. Can J Microbiol. 2012;58(1):1–9.
Cameron RK, Dixon RA, Lamb CJ. Biologically induced systemic acquired-resistance in *Arabidopsis-thaliana*. Plant J. 1994;5(5):715–25.
Carvalho TLG, Balsemao-Pires E, Saraiva RM, Ferreira PCG, Hemerly AS. Nitrogen signalling in plant interactions with associative and endophytic diazotrophic bacteria. J Exp Bot. 2014; 65(19):5631–42.
Chiarini L, Bevivino A, Dalmastri C, Nacamulli C, Tabacchioni S. Influence of plant development, cultivar and soil type on microbial colonization of maize roots. Appl Soil Ecol. 1998;8(1–3): 11–8.
Compant S, Clement C, Sessitsch A. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. Soil Biol Biochem. 2010;42(5):669–78.
da Costa PB, Beneduzi A, de Souza R, Schoenfeld R, Vargas LK, Passaglia LMP. The effects of different fertilization conditions on bacterial plant growth promoting traits: guidelines for directed bacterial prospection and testing. Plant Soil. 2013;368(1–2):267–80.
Dalmastri C, Chiarini L, Cantale C, Bevivino A, Tabacchioni S. Soil type and maize cultivar affect the genetic diversity of maize root-associated *Burkholderia cepacia* populations. Microb Ecol. 1999;38(3):273–84.
de Brito Ferreira EP, Knupp AM, Garcia Martin-Didonet CC. Growth of rice cultivars (*Oryza sativa* L.) as affected by inoculation with plant growth-promoting bacteria. Biosci J. 2014;30(3): 655–65.

de Ridder-Duine AS, Kowalchuk GA, Gunnewiek P, Smant W, van Veen JA, de Boer W. Rhizosphere bacterial community composition in natural stands of *Carex arenaria* (sand sedge) is determined by bulk soil community composition. Soil Biol Biochem. 2005;37(2):349–57.

de Souza R, Beneduzi A, Ambrosini A, da Costa PB, Meyer J, Vargas LK, et al. The effect of plant growth-promoting rhizobacteria on the growth of rice (*Oryza sativa* L.) cropped in southern Brazilian fields. Plant Soil. 2013;366(1–2):585–603.

Dobbelaere S, Croonenborghs A, Thys A, Ptacek D, Vanderleyden J, Dutto P, et al. Responses of agronomically important crops to inoculation with Azospirillum. Aust J Plant Physiol. 2001; 28(9):871–9.

Dobbelaere S, Vanderleyden J, Okon Y. Plant growth-promoting effects of diazotrophs in the rhizosphere. Crit Rev Plant Sci. 2003;22(2):107–49.

Dumbrell AJ, Ashton PD, Aziz N, Feng G, Nelson M, Dytham C, et al. Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. New Phytol. 2011;190(3):794–804.

Fang M, Kremer RJ, Motavalli PP, Davis G. Bacterial diversity in rhizospheres of nontransgenic and transgenic corn. Appl Environ Microbiol. 2005;71(7):4132–6.

Folman LB, Postma J, Van Veen JA. Ecophysiological characterization of rhizosphere bacterial communities at different root locations and plant developmental stages of cucumber grown on rockwool. Microb Ecol. 2001;42(4):586–97.

Glick BR. The enhancement of plant-growth by free-living bacteria. Can J Microbiol. 1995;41(2): 109–17.

Gopalakrishnan S, Srinivas V, Sree Vidya M, Rathore A. Plant growth-promoting activities of Streptomyces spp. in sorghum and rice. SpringerPlus. 2013;2:574.

Grayston SJ, Vaughan D, Jones D. Rhizosphere carbon flow in trees, in comparison with annual plants: the importance of root exudation and its impact on microbial activity and nutrient availability. Appl Soil Ecol. 1997;5(1):29–56.

Haichar FZ, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J, et al. Plant host habitat and root exudates shape soil bacterial community structure. ISME J. 2008;2(12):1221–30.

Haichar FZ, Santaella C, Heulin T, Achouak W. Root exudates mediated interactions belowground. Soil Biol Biochem. 2014;77:69–80.

Hardoim PR, Andreote FD, Reinhold-Hurek B, Sessitsch A, van Overbeek LS, van Elsas JD. Rice root-associated bacteria: insights into community structures across 10 cultivars. Fems Microbiol Ecol. 2011;77(1):154–64.

Hu X, Li W, Chen Q, Yang Y. Early signal transduction linking the synthesis of jasmonic acid in plant. Plant Signal Behav. 2009;4(8):696–7.

Ibekwe AM, Poss JA, Grattan SR, Grieve CM, Suarez D. Bacterial diversity in cucumber (*Cucumis sativus*) rhizosphere in response to salinity, soil pH, and boron. Soil Biol Biochem. 2010;42(4): 567–75.

Inceoglu O, Salles JF, van Elsas JD. Soil and cultivar type shape the bacterial community in the potato rhizosphere. Microb Ecol. 2012;63(2):460–70.

Jiang H, Barbier H, Brutnell T. Methods for performing crosses in *Setaria viridis*, a new model system for the grasses. J Vis Exp. 2013;80:e50527.

Kardol P, Cornips NJ, van Kempen MML, Bakx-Schotman JMT, van der Putten WH. Microbe-mediated plant-soil feedback causes historical contingency effects in plant community assembly. Ecol Monogr. 2007;77(2):147–62.

Kochian LV, Pineros MA, Hoekenga OA. The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. Plant Soil. 2005;274(1–2):175–95.

Kumar K, Muthamilarasan M, Prasad M. Reference genes for quantitative real-time PCR analysis in the model plant foxtail millet (*Setaria italica* L.) subjected to abiotic stress conditions. Plant Cell Tiss Org Cult. 2013;115(1):13–22.

Ladygina N, Hedlund K. Plant species influence microbial diversity and carbon allocation in the rhizosphere. Soil Biol Biochem. 2010;42(2):162–8.

Li P, Brutnell TP. *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. J Exp Bot. 2011;62(9):3031–7.
Li X, Rui J, Xiong J, Li J, He Z, Zhou J, et al. Functional potential of soil microbial communities in the maize rhizosphere. PLoS One. 2014;9(11):e112609.
Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core *Arabidopsis thaliana* root microbiome. Nature. 2012;488(7409):86.
Macdonald LM, Paterson E, Dawson LA, McDonald AJS. Short-term effects of defoliation on the soil microbial community associated with two contrasting *Lolium perenne* cultivars. Soil Biol Biochem. 2004;36(3):489–98.
Machado HB, Funayama S, Rigo LU, Pedrosa FO. Excretion of ammonium by *Azospirillum brasilense* mutants resistant to ethylenediamine. Can J Microbiol. 1991;57:549–53.
Mao Y, Li X, Smyth EM, Yannarell AC, Mackie RI. Enrichment of specific bacterial and eukaryotic microbes in the rhizosphere of switchgrass (*Panicum virgatum* L.) through root exudates. Environ Microbiol Rep. 2014;6(3):293–306.
Martinez-Viveros O, Jorquera MA, Crowley DE, Gajardo G, Mora ML. Mechanisms and practical considerations involved in plant growth promotion by rhizobacteria. J Soil Sci Plant Nutr. 2010;10(3):293–319.
Mazzola M, Funnell DL, Raaijmakers JM. Wheat cultivar-specific selection of 2,4-diacetylphloro glucinol-producing fluorescent Pseudomonas species from resident soil populations. Microb Ecol. 2004;48(3):338–48.
Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM, et al. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science. 2011;332(6033):1097–100.
Mitter B, Brader G, Afzal M, Compant S, Naveed M, Trognitz F, et al. Advances in elucidating beneficial interactions between plants, soil, and bacteria. Adv Agron. 2013;121:381–445.
Monteiro RA, Schmidt MA, de Baura VA, Balsanelli E, Wassem R, Yates MG, et al. Early colonization pattern of maize (*Zea mays* L. Poales, Poaceae) roots by *Herbaspirillum seropedicae* (Burkholderiales, Oxalobacteraceae). Genet Mol Biol. 2008;31(4):932–7.
Murphy JF, Zehnder GW, Schuster DJ, Sikora EJ, Polston JE, Kloepper JW. Plant growth-promoting rhizobacterial mediated protection in tomato against Tomato mottle virus. Plant Disease. 2000;84(7):779–84.
Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G, et al. Microbial diversity and microbial activity in the rhizosphere. Ciencia del Suelo. 2007;25(1):89–97.
Okon Y, Labanderagonzalez CA. Agronomic applications of Azospirillum—an evaluation of 20 years worldwide field inoculation. Soil Biol Biochem. 1994;26(12):1591–601.
Okon Y, Heytler PG, Hardy RWF. N2-fixation by *Azospirillum-brasilense* and its incorporation into host *Setaria-italica*. Appl Environ Microbiol. 1983;46(3):694–7.
Oliveira ALM, Stoffels M, Schmid M, Reis VM, Baldani JI, Hartmann A. Colonization of sugarcane plantlets by mixed inoculations with diazotrophic bacteria. Eur J Soil Biol. 2009;45(1):106–13.
Pankievicz VCS, Amaral FP, Santos FDN, Agtuca B, Xu y, Schueller MJ, Arisi ACM, Steffens MBR, de Souza EM, Pedrosa FO, Stacey G, Ferrieri RA. Robust biological nitrogen fixation in a model grass-bacterial association. Plant Journal. 2015;81, 907–919.
Pedrosa FO, Monteiro RA, Wassem R, Cruz LM, Ayub RA, Colauto NB, et al. Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. PLoS Genet. 2011;7(5):e1002064.
Pivato B, Mazurier S, Lemanceau P, Siblot S, Berta G, Mougel C, et al. Medicago species affect the community composition of arbuscular mycorrhizal fungi associated with roots. New Phytol. 2007;176(1):197–210.
Prashar P, Kapoor N, Sachdeva S. Rhizosphere: its structure, bacterial diversity and significance. Rev Environ Sci Biotechnol. 2014;13(1):63–77.
Raffi MM, Charyulu PBBN. Nitrogen fixation by the native Azospirillum spp. isolated from rhizosphere and non-rhizosphere of foxtail millet. Asian J Biol Life Sci. 2012;1:213–8.
Reinhold B, Hurek T, Niemann EG, Fendrik I. Close association of Azospirillum and diazotrophic rods with different root zones of *Kallar grass*. Appl Environ Microbiol. 1986;52(3):520–6.

Rodriguez H, Fraga R. Phosphate solubilizing bacteria and their role in plant growth promotion. Biotechnol Adv. 1999;17(4–5):319–39.
Rougier M, Chaboud A. Mucilages secreted by roots and their biological function. Israel J Bot. 1985;34(2–4):129–46.
Sage RF, Sage TL, Kocacinar F. Photorespiration and the evolution of C-4 photosynthesis. Annu Rev Plant Biol. 2012;63:19–47.
Santi C, Bogusz D, Franche C. Biological nitrogen fixation in non-legume plants. Ann Bot. 2013;111:743–67.
Schnitzer SA, Klironomos JN, HilleRisLambers J, Kinkel LL, Reich PB, Xiao K, et al. Soil microbes drive the classic plant diversity-productivity pattern. Ecology. 2011;92(2):296–303.
Siciliano SD, Theoret CM, de Freitas JR, Hucl PJ, Germida JJ. Differences in the microbial communities associated with the roots of different cultivars of canola and wheat. Can J Microbiol. 1998;44(9):844–51.
Spaepen S, Vanderleyden J. Auxin and Plant–Microbe Interactions. Cold Spring Harb Perspect Biol. 2011;3(4):a001438.
Spaepen S, Dobbelaere S, Croonenborghs A, Vanderleyden J. Effects of *Azospirillum brasilense* indole-3-acetic acid production on inoculated wheat plants. Plant Soil. 2008;312(1–2):15–23.
Steenhoudt O, Vanderleyden J. Azospirillum, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. Fems Microbiol Rev. 2000;24(4):487–506.
Tarrand JJ, Krieg NR, Dobereiner J. Taxonomic study of spirillum-lipoferum group, with descriptions of a new genus, Azospirillum gen-nov and 2 species, *Azospirillum-lipoferum* (beijerinck) comb nov and *Azospirillum-brasilense* sp-nov. Can J Microbiol. 1978;24(8):967–80.
Vargas L, Santa Brigida AB, Mota Filho JP, de Carvalho TG, Rojas CA, Vaneechoutte D, et al. Drought tolerance conferred to sugarcane by association with *Gluconacetobacter diazotrophicus*: a transcriptomic view of hormone pathways. PLoS One. 2014;9(12):e114744.
Walker V, Couillerot O, Von Felten A, Bellvert F, Jansa J, Maurhofer M, et al. Variation of secondary metabolite levels in maize seedling roots induced by inoculation with Azospirillum, Pseudomonas and Glomus consortium under field conditions. Plant Soil. 2012;356(1–2):151–63.
Wei G, Kloepper JW, Tuzun S. Induced systemic resistance to cucumber diseases and increased plant growth by plant growth-promoting rhizobacteria under field conditions. Phytopathology. 1996;86(2):221–4.
Xu J, Li Y, Ma X, Ding J, Wang K, Wang S, et al. Whole transcriptome analysis using next-generation sequencing of model species *Setaria viridis* to support C-4 photosynthesis research. Plant Mol Biol. 2013;83(1–2):77–87.
Yao H, Wu F. Soil microbial community structure in cucumber rhizosphere of different resistance cultivars to fusarium wilt. Fems Microbiol Ecol. 2010;72(3):456–63.
Zamioudis C, Pieterse CMJ. Modulation of host immunity by beneficial microbes. Mol Plant Microbe Interact. 2012;25(2):139–50.
Zehnder G, Kloepper J, Yao CB, Wei G. Induction of systemic resistance in cucumber against cucumber beetles (Coleoptera: Chrysomelidae) by plant growth-promoting rhizobacteria. J Econ Entomol. 1997;90(2):391–6.

# Chapter 15
# Herbicide Resistance in *Setaria*

**Henri Darmency, TianYu Wang, and Christophe Délye**

**Abstract** The four documented cases of field selection for herbicide resistance in weedy *Setaria* are described in this chapter. In each case, weed control failure was observed in practice in the field. In all cases, resistance was target-site-based resistance and was due to single nucleotide mutations causing amino-acid substitutions at codon 264 of psbA (photosystem II inhibitors), codons 136 and 239 of α2-tubulin (tubulin polymerization inhibitors), codon 1781 of acetyl-CoA carboxylase (acetyl-CoA carboxylase inhibitors), or codons 653 or 654 of acetolactate-synthase (acetolactate-synthase inhibitors). The heredity of resistance in these cases was maternal, nuclear recessive, nuclear dominant, or partially dominant, respectively. Pleiotropic effects of the mutant alleles were observed on seed production for the herbicide-resistant alleles Gly-264 of psbA and Ile-239 of α2-tubulin (22 % yield reduction for both alleles), but not for the Leu-1781 acetyl-CoA carboxylase allele. These alleles were introgressed in foxtail millet (*S. italica*) to develop herbicide-resistant genetic resources and germplasm with the aim to produce and release elite varieties of foxtail millet. This material was also used to study pollen dispersal and possible gene flow between weedy *Setaria* and cultivated foxtail millet.

**Keywords** *Setaria* • Millet • Foxtail • Weed • Herbicide • Resistance • Gene flow

## 15.1 Introduction

From the middle of the nineties, weed control in arable fields, roadsides, urban and industrial areas has most often relied upon herbicide spray. In various places where the same herbicide was continuously used, herbicide-resistant plants were selected and have caused trouble to farmers (Beckie and Tardif 2012; Délye et al. 2013). Globally, there are 245 species that have evolved resistance to 22 of the 25

H. Darmency, M.Sc., Ph.D. (✉) • C. Délye
Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne Franche-Comté,
F-21000 Dijon, France
e-mail: henri.darmency@dijon.inra.fr

T. Wang
Institute of Crop Sciences, Chinese Academy of Agriculture Sciences, Beijing, 100081 China



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_15

known herbicide modes of action, in 85 crops and in 66 countries (Heap 2015). Different mechanisms are responsible for resistance, an adaptive response of weeds to the herbicide selection pressure: (1) escape of the spraying period via modified phenology; (2) reduction in herbicide penetration through modified cuticle properties; (3) altered translocation of the herbicide toward its target site; (4) sequestration of the herbicide away from its target site; (5) enhanced degradation of the herbicide; (6) mutation at the herbicide target site; (7) overproduction of the herbicide target site; and (8) compensation for deleterious effects of the action of the herbicide (Délye et al. 2013).

The *Setaria* genus also demonstrates herbicide resistance. While one species (*S. italica*, foxtail millet) is a staple crop in Asia and Africa, the most widespread *Setaria* species are serious arable weeds, possibly because their original natural habitats became cultivated or managed by human activities after the onset of agriculture: species adapted to highly disturbed arable fields are indeed offered vast surfaces as potential habitat (Dekker 2004). Weedy *Setaria* include *S. viridis* (L.) P. Beauv. (green foxtail), *S. verticillata* (L.) P. Beauv. (bristly foxtail), *S. faberi* F. Hermann (giant foxtail), *S. pumila* (Poiret) Roemer & Schultes (syn. *S. glauca*) (yellow foxtail), and *S. parviflora* (Poir.) Kerguélen S. (syn. *S. geniculata*) (knotroot foxtail). As weeds, these species have been subjected to repeated selection by herbicides over broad areas and consequently have evolved herbicide resistance. Understanding the evolution of herbicide resistance is of scientific relevance because of the impact of this resistance in agriculture and because research addressing herbicide resistance mechanisms and evolution allows considerable insight into plant physiology and response to selection. In the first part, we review herbicide resistance cases reported in *Setaria* species, and their underlying genetics, mechanisms, and biological consequences. We analyze in particular the fitness cost, estimated by comparing resistant and susceptible material sharing a common origin (see Vila-Aiub et al. 2009 for a review). In the second part, we summarize the efforts implemented to transfer genes responsible for herbicide resistance to cultivated varieties of *Setaria* in order to facilitate weed control in those crops. Indeed, developing herbicide-resistant cultivars in "minor" crops could be a way to maintain the global diversity of cropping systems. The alternative would be laying aside these crops because they are often not considered profitable enough by agrochemical companies to foster the development of selective herbicides, and the cost and time required for mechanical or hand weeding render them unattractive to growers. As far as we know, introgressing herbicide resistance is one of the rare cases of using wild *Setaria* in foxtail millet breeding programs.

## 15.2 Herbicide Resistance in Weedy *Setaria*

The high variability described for the *Setaria* genus prompted early researchers in the field of herbicide-based weed control to investigate natural variation in sensitivity to herbicides. Variation in efficacy of the herbicide dalapon, a lipid synthesis inhibitor (Herbicide Resistance Action Committee (HRAC) group N), was observed

among various accessions of *S. pumila* and *S. faberi*, which could explain the reported variable efficacy in controlling these species in the fields (Santelmann and Meade 1961). Although they observed a lack of control for some populations that eventually produced seeds, the authors did not use the term "resistance" but referred to "variation of dalapon susceptibility." Similarly, Oliver and Schreiber (1971) observed a differential efficacy of the two photosynthesis inhibitors herbicides atrazine and propazine (HRAC group C1) among *S. viridis* forms, including various forms corresponding to spp. *pycnocoma* (Steudel) Tzvelev. A relationship between variation in the capacity to metabolize herbicides and variation in sensitivity observed among the different *S. viridis* forms was subsequently established (Thompson 1972). Small (two- or three-fold) intraspecific differences in sensitivity to atrazine among populations were subsequently confirmed in *S. viridis*, *S. adhaerens*, *S. verticillata*, and *S. pumila* (De Prado et al. 1990; Wang and Dekker 1995). Similar resistance ratios were also observed for metolachlore, a cell division inhibitor (HRAC group K3) in *S. viridis* and *S. pumila* (Wang and Dekker 1995). Intraspecific variation in herbicide detoxification could be at the root of these differences. However, all these cases of variation in sensitivity to herbicides were not related to a documented specific and repeated herbicide use of the herbicide and did not lead to weed control failure. As such, they did not fall within the definition of herbicide resistance in weeds (Heap 2015) and were considered representing the standing variation in sensitivity of the *Setaria* species. The four documented cases of resistance in *Setaria* are reported below in chronological order (Table 15.1). Rapid biological tests at the seed germination and seedling stage, as well as molecular tools, were set up to identify the resistant mutants (Fig. 15.1). Typical evolution of herbicide resistance had also been observed in response to a long and repeated use of acetochlor, a cell division inhibitor (HRAC group K3), but this case was not further investigated (Baeva 2007).

**Table 15.1** Summary of the characteristics of the four documented herbicide-resistance cases in *Setaria viridis* and date of the first field record

| Herbicide | Date | HRAC group | S/R plant | S/R target site | Field rate | Inheritance | Gene | Codon substitution | Fitness |
|---|---|---|---|---|---|---|---|---|---|
| Atrazine | 1980 | C1 | >50 | 1000 | 5× | Maternal | psbA | Ser264-Gly | −22 % |
| Trifluralin | 1987 | K1 | 7 | ND | 0.6× | Nuclear, recessive | α2-tubulin | Leu136-Phe | −20 % |
| | | | | | | | | Thr239-Ile | |
| Sethoxydim | 1990 | A | 2980 | 700 | >2× | Nuclear, dominant | ACCase | Ile1781-Leu | = |
| Imazethapyr | 2001 | B | 182 | 260 | >2× | Nuclear, ND | ALS | Ser653-Thr/Asn/Ile | ND |
| | | | | | | | | Gly654-Asp | |

The R/S resistance factors (R/S ratio of the herbicide doses which cause 50 % mortality of a plant population or 50 % inhibition of growth or other vital physiological function) show the values expressed at the whole-plant level and at the target site (chloroplast or enzyme activity) recorded for the most resistant accession
*ND* not determined

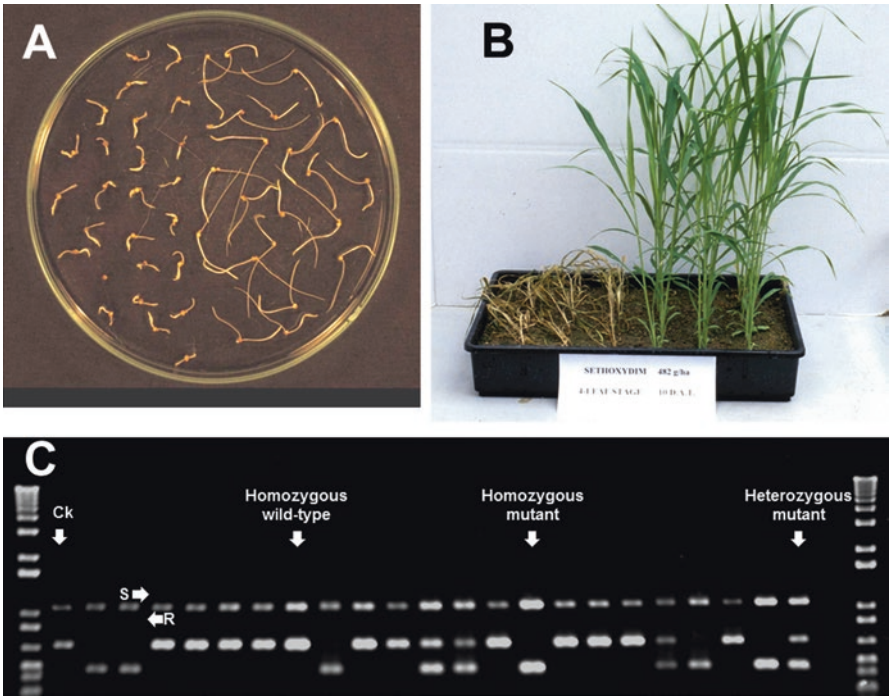


**Fig. 15.1** Examples of identification of herbicide-resistant plants using different types of assays. (**a**) Petri dish assay showing sensitive (*left*) and resistant (*right*) seedlings growing on a medium containing a tubulin polymerization inhibitor (HRAC group K1). Growth of sensitive seedlings is reduced with distorted shoot and roots, while growth of resistant seedling is unaffected. (**b**) Whole-plant spraying assay showing sensitive (*left*, killed) and resistant (*right*, unaffected) seedlings 10 days after application of a commercial formulation of an acetyl-CoA carboxylase inhibitor (HRAC group A) at the 3–4 leaf stage. (**c**) Genotyping assay showing the detection of Leu-1781 ACCase mutant alleles using allele-specific PCR as described in Délye et al. (2002) (Ck, internal positive control (1087 bp); S, amplicon specific for ACCase alleles carrying a wild-type 1781 codon (Ile-1781, 677 bp); R, amplicon specific for ACCase alleles carrying an herbicide-resistant, mutant 1781 codon (Leu-1781, 448 bp) [(**c**) reproduced with permission of Springer]

### *15.2.1 Resistance to Photosystem II Inhibitors*

The triazine herbicide family (HRAC group C1) inhibits electron transfer at the photosystem II (PSII) level. Triazines were massively and repeatedly used to control weeds in maize monoculture during the seventies. Although *Setaria* species display some detoxification capacity against triazines (Gimenez-Espinosa et al. 1996), triazine applications generally resulted in a nearly total control of these weeds. The first case of resistance to triazines in *Setaria* was observed in a maize monoculture in France, where plants surviving three-fold the atrazine field rate evolved in a *S. viridis* population that had been sprayed with this herbicide for 7 years (Gasquez and Compoint 1981). Similar resistance evolution was subsequently observed in

Spain (De Prado et al. 2000) and Yugoslavia (Konstantinovic 2001) in *S. viridis* and in France in *S. viridis* spp. *pycnocoma* (Darmency and Pernès 1985). Other *Setaria* species were also involved: *S. pumila* in eastern Canada and Spain (Stepenson et al. 1990; De Prado et al. 1989), and *S. faberi* in the USA and Spain (Ritter et al. 1989; De Prado et al. 2000). All these confirmed resistance cases evolved in fields grown with maize and with a long history of atrazine applications. Resistant *S. viridis* have also been found in French vineyards continuously treated with triazine herbicides (Darmency, unpublished). Resistant plants withstood up to ten times the herbicide field dose, a rate at which the maize crop is killed. Further investigation showed that photosynthetic electron transport was unaffected by the herbicide in the resistant plants, with a resistance factor (i.e., R/S ratio of the herbicide doses which cause 50 % mortality of a plant population or 50 % inhibition of growth or other vital physiological function) ranging from 300 to 1000 as evaluated at the chloroplast activity level. A rapid and simple fluorescence test allowed easy resistance diagnosis (Gasquez and Compoint 1981; Ritter et al. 1989; De Prado et al. 1989, 2000).

Today, resistance to triazines has been reported in 72 weed species (Heap 2015). Resistant plants showed target site-based resistance resulting from a mutation at the herbicide-binding site, a chloroplast 32 kDa polypeptide called D1 and encoded by the chloroplastic gene psbA (see Tian and Darmency 2006 for a review). In most cases, this modification was a Ser-to-Gly substitution at amino-acid residue 264 protein D1. This results in an altered conformation of the herbicide-binding site on protein D1 causing a drastic reduction in herbicide binding. The Gly264 psbA allele was identified in the French accessions of *S. viridis* exhibiting resistance to triazines (Tian and Darmency 2006; Jia et al. 2007). Being chloroplast encoded, triazine resistance is expected to be maternally inherited. This was confirmed by interspecific crosses between the sexually compatible and closely related species *S. viridis* and foxtail millet (*S. italica*): only hybrid progeny derived from *S. viridis* resistant mother-plants inherited resistance, and resistance did not segregate in the $F_2$ generation (Darmency and Pernès 1985). However, maternal inheritance was not absolute. Analysis of >750,000 hybrid plants produced using a male sterile foxtail millet variety crossed with triazine-resistant *S. viridis* and confirmed as hybrids by reciprocal markers revealed pollen-mediated transfer of the chloroplast resistance gene: the sensitive female parent produced 0.03 % resistant progeny.

The herbicide-resistant Gly264 psbA allele entails physiological consequences: a less efficient electron transport through the PSII and a series of functional and anatomical alterations of the chloroplast, which cause a strong fitness cost (Arntz et al. 2000). Fitness cost was not directly investigated in weedy *Setaria*. Analysis of a series of backcrossed progeny derived from a cross between triazine-resistant *S. viridis* carrying Gly264 psbA and foxtail millet showed that resistant $BC_2$ progeny had a lower rate of photosynthesis ($CO_2$ fixation) than their sensitive counterparts (i.e., the reciprocal $BC_2$) at 27 °C, a normal growth temperature for this summer growing plant. No difference was observed at lower temperatures (Ricroch et al. 1987). This could be due to a chlorophyll a/chlorophyll b ratio that was 10 % lower in the resistant *S. viridis* plants than in the sensitive plants

(Darmency et al. 1992). In field experiments, seed production of the resistant *S. viridis* plants was 22 % lower than that of the sensitive counterparts (Darmency and Pernès 1989). In greenhouse experiments, high plant density conditions further decreased the relative productivity of the resistant *S. viridis* plants up to 65 % (Reboud and Till-Bottraud 1991).

### *15.2.2 Resistance to Tubulin Polymerization Inhibitors*

The dinitroaniline herbicide family (HRAC group K1) inhibits cell division. At the end of the eighties, resistance to dinitroaniline herbicides evolved in several populations of *S. viridis* in the Canadian prairies that received at least 3–5 applications of these herbicides in 10 years (Morrison et al. 1989). *S. viridis* was the only *Setaria* species reported to have evolved a resistance to dinitroaniline leading to practical control failure. A rapid Petri dish bioassay allowed clear discrimination of resistant and sensitive young *S. viridis* seedlings on the basis of inhibition of radicle growth (Beckie et al. 1990). Segregation studies identified a 3:1/sensitive:resistant ratio in the $F_2$ generation derived from crosses between dinitroaniline-resistant and sensitive *S. viridis* plants, showing that the resistance was under the control of a recessive nuclear locus (Jasieniuk et al. 1994). Resistance factors were moderate (c.a. seven-fold), but the resistant plants were cross-resistant to all dinitroaniline herbicides (Beckie and Morrison 1993a, b) as well as to other unrelated tubulin-destabilizing drugs. This suggested that the resistant plants may contain an altered protein that stabilizes microtubule formation (Smeda et al. 1992). Tubulins are dimeric proteins consisting into one α and one β subunit that polymerize into microtubules (Breviario et al. 2013). Four tubulin genes (two α and two β) were identified in *S. viridis* (Délye et al. 2004). A Leu136-Phe or a Thr239-Ile substitution in the gene encoding the α2-tubulin was found responsible for resistance (Délye et al. 2004). Occurrence of two mutant α2-tubulin alleles was necessary to confer resistance, confirming the recessive status of this resistance (Jasieniuk et al. 1994). Tridimensional modelling shed light on the stereochemical organization of the α2-tubulin region involved in herbicide binding and tubulin polymerization (Délye et al. 2004). Allele-specific polymerase chain reaction (PCR) assays were set up to allow quick discrimination of the different α2-tubulin alleles (Délye et al. 2004, 2005).

Resistance to dinitroanilines was the first demonstrated case of recessive control of a resistance to herbicides in weeds. Recessive control is not favorable to the establishment of a resistance in the field. However, *S. viridis* is highly self-fertilized and produces huge number of seeds, a feature facilitating the emergence of homozygous resistant plants. The frequency of dinitroaniline-resistant plants did not vary after 7 years with no dinitroaniline application, suggesting there was no substantial fitness penalty associated to this resistance (Andrews and Morrison 1997). However, the persistence of resistance could also be due to the long-term persistence of the herbicide in the soil and/or to a large resistant soil seed bank established during the

years *S. viridis* was not controlled. Indeed, comparison of dinitroaniline-resistant (Ile239 allele) and sensitive nearly isogenic *Setaria* material identified a reduction in seed production of about 20 % in the resistant lines (Darmency et al. 2011). This fitness cost was similar to that found for resistance to triazines and was confirmed by field experiments, where Ile239-dinitroaniline- and triazine-resistant lines tested together showed similar seed production (Wang et al. 2010a).

### *15.2.3 Resistance to Acetyl-CoA Carboxylase (ACCase) Inhibitors*

The herbicides inhibiting ACCase (HRAC group A) specifically disrupt fatty acid biosynthesis in the Gramineae. Resistance to ACCase inhibitors evolved in the early nineties in several populations of *S. viridis* (Heap and Morrison 1996) and *S. faberi* (Stoltenberg and Wiederholt 1995) following the repeated (around seven times) use of these herbicides during one decade. Resistance increased rapidly in *S. viridis* in Canada. It was detected in 6 % of the fields sampled in a large survey in 2001–2003, and in 27 % of those sampled in 2007–2011 (Beckie et al. 2013). It was later found in Spain (De Prado et al. 2004). The resistant *Setaria* plants showed cross-resistance to the majority of ACCase inhibitors with various resistance factors, with particularly high resistance levels to the herbicide sethoxydim (resistance factors ranging from 20 to 2900: Heap and Morrison 1996). The ACCase enzyme extracted from resistant *S. viridis* or *S. faberi* plants was much less sensitive to ACCase inhibitors than that from sensitive plants (Marles et al. 1993; Shukla et al. 1997; Volenberg and Stoltenberg 2002a; De Prado et al. 2004). Subsequent segregation studies using hybrid progenies obtained by crossing *S. viridis* and foxtail millet showed that resistance to the ACCase inhibitor sethoxydim was due to a single, completely dominant, nuclear locus (Wang and Darmency 1997b). Similarly, a single, nuclear, co-dominant locus controlled the response to the ACCase inhibitor fluazifop in *S. faberi* (Volenberg and Stoltenberg 2002b).

A point mutation causing an Ile1781-Leu substitution in the carboxyltransferase domain of the nuclear gene encoding the plastidic ACCase isoform was demonstrated to be a cause for resistance to the ACCase inhibitors sethoxydim resistance in *S. viridis* (Délye et al. 2002). Other mutations that confer different patterns of cross-resistance have subsequently been identified in other grass weed species (Délye 2005; Beckie and Tardif 2012). Observation of different patterns of cross-resistance among *S. viridis* populations (Heap and Morrison 1996; Beckie et al. 1999) most likely denotes the presence of mutant, herbicide-resistant ACCase alleles different from the Leu1781 allele or from other resistance mechanisms in this species. Complete proof that the Leu1781 allele does encode for resistance was provided by the transfer of the mutated gene sequence to maize through genetic engineering resulting in herbicide-resistance expression (Dong et al. 2011).

Using nearly isogenic plant material derived from interspecific crosses between *S. viridis* and foxtail millet, more vigorous juvenile growth in the field,

earlier flowering, a higher number of tillers and grains were recorded on the resistant plants carrying Leu1781 ACCase allele than on their sensitive counterparts (Wang et al. 2010b). The differences were exacerbated when both genotypes were grown in mixture. The seeds of the Leu1781 ACCase plants were lighter than those of the sensitive plants although more abundant (Wang et al. 2010b). Fitness of both genotypes over the whole life cycle was not different in a 3-year experiment when plots of mixed populations were left unmanaged, but an excess of Leu1781 ACCase plants was found in plots where low doses of trifluralin herbicide (HRAC group K1, not related to the mode of action of ACCase inhibitors and ACCase resistance) created stressing conditions (Wang et al. 2010b). Therefore, fitness neutrality or benefit of the Leu1781 ACCase allele (or of a closely linked gene) exists and was triggered by the habitat conditions (Wang et al. 2010b). This predicts a long and successful persistence of plants carrying Leu1781 ACCase in the field.

### *15.2.4 Resistance to Acetolactate-Synthase (ALS) Inhibitors*

The herbicides inhibiting ALS (HRAC group B) disrupt the biosynthesis of branched-chain amino acids. Resistance to these herbicides evolved in several populations of *S. viridis*, *S. pumila*, and *S. faberi* in Canada and the USA at the end of the nineties after one or two yearly applications of these herbicides during 4–9 years (Volenberg et al. 2001, 2002; Heap 2015). Dose-response experiments on whole *S. faberi* plants showed resistance factors of about 10–20 that varied according to population and the herbicide (Volenberg et al. 2001). In vitro ALS enzyme assay showed that the enzyme of resistant *S. faberi* plants was resistant to herbicides. Resistance segregation in *S. faberi* showed a 1:2:1 sensitive:intermediate:resistant segregation in $F_2$, thus indicating that resistance was controlled by a nuclear semidominant locus (Volenberg et al. 2001). Since *S. faberi* is an allotetraploid (Benabdelmouna et al. 2001), it is likely that only one locus is involved here with simple disomic inheritance. Similar results were obtained for *S. viridis*, with resistance factors varying with the herbicide tested (Volenberg et al. 2002).

In other *S. viridis* populations, resistance was demonstrated to be due to mutant alleles encoding ALS enzymes carrying amino-acid substitutions that modified the herbicide-binding site (Laplante et al. 2009). Different substitutions were identified: Ser653-Thr, Ser653-Asn, Ser653-Ile and Gly654-Asp. They conferred different cross-resistance patterns to ALS inhibitors (Laplante et al. 2009). Recently, the first case of *S. viridis* with nontarget-site resistance has been identified for a population in maize in France (Délye, unpublished).

No report has been published to date for *Setaria* species on the possible consequences on plant fitness of one mutant ALS allele. A few indications may be inferred from "herbicide-tolerant" crop cultivars carrying natural or induced similar mutations: an Asn653 ALS allele is present in maize, rice, oilseed rape and wheat culti-

vars of the brand Clearfield® and no any deleterious effect has been recorded on yield; similarly, there was no effect of this mutation on seed production in *Arabidopsis* (see Darmency 2013 for review).

## 15.3 Herbicide-Resistant Crop Varieties of Foxtail Millet (*S. italica*)

Breeding for herbicide resistance (sometimes referred to as "tolerance") is a very recent trend in plant breeding. Demonstrating the single-gene control of resistance to triazines, i.e., a single gene controlling a drastic change in phenotype, was certainly a major incentive to this approach: breeders realized that a single gene could make herbicides selective for hitherto sensitive crop varieties, as illustrated for triazine-resistant oilseed rape (Beversdorf et al. 1980). This approach is particularly attractive considering the current lack of new herbicides released by the industry (Duke 2012). In addition, few herbicides selective for foxtail millet are marketed because this crop does not represent a market profitable enough to justify the expenses pertaining to herbicide development and commercial release. Accordingly, weed control remains one of the major issues when growing foxtail millet (Shanxi Academy of Agricultural Sciences 1987; Zhou et al. 2013). Herbicide-resistant foxtail millet cultivars were thus bred to overcome this situation. Since no herbicide-resistant germplasm was available for foxtail millet (Shanxi Academy of Agricultural Sciences 1987), herbicide-resistant genotypes of the sexually compatible and closely related weed *S. viridis* described in the preceding sections were used as a source for the resistance trait. At the same time, breeding herbicide-resistant foxtail millet aroused considerable concern about the potential dispersal of the genes endowing herbicide resistance back to weedy *Setaria* populations.

### *15.3.1 Introgression of Herbicide Resistance Genes*

Triazine-resistant foxtail millet germplasm was first generated from an interspecific cross with *S. viridis* with the weed as the female in order to retain the chloroplast encoded Gly264 psbA allele conferring triazine resistance (Darmency and Pernès 1985). Two backcross generations combined with morphological selection were enough to eliminate weedy traits and generate resistant foxtail millet germplasm (Naciri et al. 1992). High-yield foxtail millet lines were ultimately derived from these crosses (Ji et al. 2006; Shi et al. 2008) despite the fitness cost associated to Gly264 psbA.

Dinitroaniline-resistant germplasm was then generated. The recessive nature of resistance associated with the causal Ile239 α2-tubulin allele complicated the breeding scheme. In addition, segregation distortion against homozygous mutant hybrid

progeny was observed with on average 15 % homozygous resistant progeny plants instead of the expected 25 % (Wang et al. 1996b). The segregation distortion was attributed to linkage of the Ile239 α2-tubulin allele with a modifier gene whose expression was only observed in the interspecific hybrids (Tian et al. 2006). It was inferred from alignment of the rice and millet genetic maps that the α2-tubulin gene belongs on linkage group IX, a chromosome that showed distorted segregation in an independent RFLP study (Wang et al. 1998). The linkage of the resistance gene with the putative gametophyte gene resulting in 69 % gamete viability could be broken in advanced backcross progeny (Tian et al. 2006). Resistant foxtail millet lines were further selected and a resistance pattern was observed that was similar to that observed for the original *S. viridis* genotype (Wang and Darmency 1997a). As the resistance factor obtained in the foxtail millet germplasm was not high enough to enable fully satisfactory weed control in the field, the Ile239 α2-tubulin foxtail millet germplasm was not used for commercial release, but it was helpful to facilitate hybrid seed production as described below (Wang et al. 1996a).

Foxtail millet germplasm obtained by introgressing Leu1781 ACCase from *S. viridis* was used in a breeding program intended to release a series of commercial, herbicide-resistant cultivars. The resistance pattern observed in the *S. viridis* parent was transferred to the derived foxtail millet lines (Wang and Darmency 1998). Before the point mutation was elucidated, AFLP markers were developed to identify the mutation throughout the breeding scheme (Niu et al. 2002). The higher seed production associated to the Leu1781 ACCase allele is a desired herbicide-resistant resource for breeding, and was combined with a restorer for seed size in further crosses (Wang et al. 2000, 2010c).

Based on this research, Wang's team at the Chinese Academy of Agricultural Science established a nation-wide cooperation network in China for the breeding of herbicide-resistant varieties in foxtail millet. A cooperative group distributed new and improved herbicide-resistant materials and breeding techniques to local breeders to help further advance adoption of varieties. Millet breeders in different ecological areas used existing commercial varieties that exhibited strong performance as the recurrent parent to improve herbicide-resistant materials for high yield and high herbicide resistance, and to generate lines that showed an aggregation of desirable traits. These methodologies effectively accelerated the breeding process of new varieties and hybrids, and their application to the field. There are now presently 30 novel herbicide-resistant millet varieties/hybrid varieties registered at the national or local level in China. These new varieties are now being widely used in all three major millet-producing regions of China. Resistant varieties employed in the Mid-northern region of China include SR3522, Jigu 24, Jigu 25, Jigu 29, Changgu 2, and zhangzagu 3, those employed in the West-northern region include zhangzagu 3, zhangzagu 5, zhangzagu 6, zhangzagu 9, Bagu 214, and Longgu 11, and the East-northern region relies heavily on zhangzagu 3 zhangzagu 5, Chizagu 1, and Jigu 24. Since this is a new technology, it is necessary to use a new way of promotion. Zhangjiakou Academy of Agricultural Sciences and allied teams established a novel system that connects millet breeding, planting, processing and marketing, which has achieved considerable positive social and economic influence (Li et al. 2014; Song et al.

2014). To date, all new herbicide-resistant millet varieties have demonstrated characteristics such as efficient herbicide resistance, high and stable yield and strong adaptability. Since 2001, the newly-bred varieties/hybrids have been grown on more than 0.8 million hectares (customarily, foxtail millet acreage in China is 1–1.5 million hectares each year). Some varieties were also trialled in a large area in Ethiopia and other Africa countries (Liu and Zhao 2012; Hao et al. 2013).

During the process, different types of herbicide resistance have been developed to secure the selection and production of hybrid varieties expressing heterosis and potentially having higher yield (Wang et al. 2000). The resistance traits to triazine (cytoplasmic inheritance) and dinitroaniline herbicides (nuclear recessive) were transferred to male sterility lines. Using these triazine-resistant male sterility lines and corresponding herbicides it is possible to simplify seed production procedures and improve seed production yield. In addition, the use of dinitroaniline-resistant male sterility lines would help to maintain line purity and to reduce outcrossing so that there would be less need to isolate the reproduction field. The resistance traits to herbicides inhibiting ACCase (nuclear dominant) combined with a restorer provides a mechanism to efficiently eliminate false positive hybrids (mixed mother parent seedling and weeds) (Wang et al. 2000, 2010c; Tian et al. 2010). The successful breeding of various herbicide-resistant millet varieties has been instrumental in pursuing these objectives.

### *15.3.2 Gene Flow*

Although *Setaria* species are primarily self-pollinated, pollen can move several dozen meters from the source plants and fertilize male-sterile as well as male-fertile plants (Wang et al. 1998, 2001). *S. viridis* ssp. *pycnocoma* is considered to be the result of ancient hybridization between *S. viridis* and foxtail millet (Darmency 2004). These species constitute a dynamic and evolving "weed-crop complex" (Darmency 2004), and spontaneous crosses could have occurred reciprocally (Till-Bottraud et al. 1992). Thus, when *S. viridis* grows close to foxtail millet, gene flow is unavoidable. Hybridization of other weedy *Setaria* species with foxtail millet is far less likely (Darmency and Dekker 2011). Average outcrossing rate for *S. viridis* plants planted 0.25 m apart was 0.74 % (Till-Bottraud et al. 1992), and 0.48 % for *S. faberi* plants planted 0.36 m apart (Volenberg and Stoltenberg 2002b). Around 0.2 % interspecific hybrids were produced by *S. viridis* plants because of pollination by foxtail millet in field (row) conditions (Darmency et al. 1987; Till-Bottraud et al. 1992), and up to 3 % hybrids were recorded when plants of the two species were grown in close mixture (De Wet et al. 1979). Under commercial field conditions, the rate of hybrid produced by *S. viridis* pollinated by foxtail millet was lower, ranging from 0.039 % within the foxtail millet field to 0.002 % 20 m from the field (Shi et al. 2008). After 6 years of testing, results show that the gene flow from *cv.* to wild population effectively occurs in production condition but is manageable when the herbicide selection stops so that it can be controlled through crop and herbicide rotation.

*In Natura*, hybrids of *S. viridis* and foxtail millet are expected to suffer a fitness penalty due to a mix of antagonistic wild and domesticated characters (i.e., flowering synchrony, seed shedding, seed size, seed dormancy). Although no direct estimate of the relative fitness cost of hybridization has ever been carried out, we indirectly calculated from our own experiments that $F_1$ hybrids of *S. viridis* and foxtail millet produces 15–30 times less viable seeds than a *S. viridis* plant (Darmency, unpublished). However, in the subsequent generations, seed fertility is rapidly restored by back-cross with *S. viridis*, although the number of tillers remains low compared to *S. viridis*. In field experiments with a foxtail millet cultivar carrying the dominant Leu1781-ACCasa allele, only a slow increase in herbicide-resistant progeny ($F_1$ and hybrid descendants of *S. viridis*) was observed during 4 years of herbicide-free cultivation of the resistant cultivar (0.1 % after 4 years) (Shi et al. 2008). This proportion decreased rapidly in the absence of the resistant foxtail millet cultivar to reach 0.01 % within 2 years (Shi et al. 2008). Using herbicides to which resistance has been introgressed in foxtail millet is obviously expected to facilitate the selection of the resistant progeny of the hybrids, which would jeopardize the herbicide-resistant cultivar strategy. For this strategy to be efficient, it is clearly necessary to closely monitor the increase in frequency of the resistance genes in weedy *Setaria* populations and, if possible, to alternate growing foxtail millet cultivars with resistance to different herbicide modes of action in a given field.

## 15.4 Perspectives

Although weedy *Setaria* are widespread weeds, few *Setaria* populations evolved herbicide-resistance in comparison to other grass weed genera (e.g., *Lolium, Alopecurus, Echinochloa, Poa*: Heap 2015). Only four herbicide modes of action are affected by resistance in *Setaria* species. Non-target-site-resistance that is a major cause for resistance in other grasses (Beckie and Tardif 2012; Délye et al. 2013) has not been identified to date in *Setaria*, although some studies identified the potential for this non-target-site based resistance to evolve in *Setaria* (Santelmann and Meade 1961; Oliver and Schreiber 1971; Thompson 1972; De Prado et al. 1990; Wang and Dekker 1995). This situation may be due to non-target-site based resistance having been overlooked by researchers that were more focussed on target-site-based resistance. Non-target-site based resistance is largely considered to evolve by accumulation of genes with additive effects in a same plant via sexual reproduction (Délye et al. 2013). The strong autogamy of the *Setaria* species may thus also be a reason for the absence of report of non-target-site based resistance in these species. However, multiple resistance to ACCase and microtubule inhibitors was detected in some locations in Canada, which may confirm the potential for further evolution in response to environmental conditions (Beckie et al. 1999). In contrast to *S. viridis* and *S. faberi* for which resistance cases have been reported, there is an absence of reported resistance cases in *S. adhaerens* and *S. verticillata,* which may be due to moderate herbicide use in the distribution areas of these species, i.e., warmer and more tropical

zones. Genome-based differences also could contribute to this difference because *S. viridis* (diploid genome A) and *S. faberi* (tetraploid genomes A and B) share in common the A genome while *S. adhaerens* has the B genome. However, both *S. faberi* and *S. verticillata* are allotetraploid and carry genomes A and B, thus casting some doubt on a genome-mediated effect. Introgression of the resistance genes into foxtail millet has proven to be an efficient strategy to generate herbicide-resistant cultivars. However, such cultivars must be used with care because of the high risk for the transfer by gene flow of the herbicide-resistant allele back to weedy *Setaria* species. We have already paid attention to this topic, especially in view of releasing future genetically engineered lines into production. Fortunately, we have not found the problem in the fields of foxtail millet, as well as in other crops up to now. Perhaps this is due to the fact that above-mentioned herbicides, especially ACCase inhibitors, are not utilized in field production so much. In addition, it could be due also to field scouting and hand weed control of remaining weeds since Chinese farmers usually deal with small-scale fields by hand. In any case, a more in-depth understanding of the genetic relationships between *S. viridis* and foxtail millet may be necessary to be able to correctly assess the risk for interspecific herbicide resistance gene flow at both field and landscape levels and to devise more appropriate recommendations for the use of herbicide-resistant foxtail millet cultivars.

## References

Andrews TS, Morrison IN. The persistence of trifluralin resistance in green foxtail (Setaria viridis) populations. Weed Technol. 1997;11:369–72.

Arntz AM, De Lucia EH, Jordan N. From fluorescence to fitness: variation in photosynthetic rate affects fecundity and survivorship. Ecology. 2000;81:2567–76.

Baeva G. Resistance study of Setaria glauca PB to the herbicide acetochlor. In: Proceedings of the 14th EWRS symposium, Hamar, Norway; 17–21 June 2007. p. 146.

Beckie HJ, Morrison IN. Effect of ethalfluralin and other herbicides on trifluralin-resistant green foxtail (Setaria viridis). Weed Technol. 1993a;7:6–14.

Beckie HJ, Morrison IN. Effective kill of trifluralin-susceptible and -resistant green foxtail (Setaria viridis). Weed Technol. 1993b;7:15–22.

Beckie HJ, Tardif FJ. Herbicide cross resistance in weeds. Crop Prot. 2012;35:15–28.

Beckie HJ, Friesen LF, Nawolsky KM, Morrison IN. A rapid bioassay to detect trifluralin-resistant green foxtail (Setaria viridis). Weed Technol. 1990;4:505–8.

Beckie HJ, Thomas AG, Légère A. Nature, occurrence, and cost of herbicide-resistant green foxtail (Setaria viridis) across Saskatchewan ecoregions. Weed Technol. 1999;13:626–31.

Beckie HJ, Lozinski C, Shirriff S, Brenzil CA. Herbicide-resistant weeds in the Canadian Prairies: 2007 to 2011. Weed Technol. 2013;27:171–83.

Benabdelmouna A, Shi Y, Abirached-Darmency M, Darmency H. Genomic in situ hybridization (GISH) discriminates between the A and B genomes in diploid and tetraploid Setaria species. Genome. 2001;44:685–90.

Beversdorf WD, Weiss-Lerman J, Reickson LR, Souza-Machado V. Transfer of cytoplasmically-inherited triazine resistance from bird's rape to cultivated rapeseed (Brassica campestris L. and B. napus L.). Can J Genet Cytol. 1980;22:167–72.

Breviario D, Giani S, Morello L. Multiple tubulins: evolutionary aspects and biological implications. Plant J. 2013;75:202–18.

Darmency H. Incestuous relations of foxtail millet (Setaria italica) with its parents and cousins. In: Gressel J, editor. Crop ferality and volunteerism: a threat to food security in the transgenic era. Boca Raton: CRC Press; 2004. p. 81–96.
Darmency H. Pleiotropic effects of herbicide-resistance genes on crop yield: a review. Pest Manag Sci. 2013;69:897–904.
Darmency H, Dekker J. Setaria. In: Kole C, editor. Wild crop relatives: genomic and breeding resources. Millet and grasses. Berlin: Springer; 2011. p. 275–96.
Darmency H, Pernès J. Use of wild Setaria viridis (L) Beauv to improve triazine resistance in cultivated S. italica (L) by hybridization. Weed Res. 1985;25:175–9.
Darmency H, Pernès J. Agronomic performance of a triazine resistant foxtail millet (Setaria italica (L) Beauv). Weed Res. 1989;29:147–50.
Darmency H, Zangre GR, Pernès J. The wild-weed-crop complex in Setaria: a hybridization study. Genetica. 1987;75:103–7.
Darmency H, Chauvel B, Gasquez J, Matejicek A. Variation of chlorophyll a/b ratio in relation to population polymorphism and mutation of triazine-resistance. Plant Physiol Biochem. 1992;30: 57–63.
Darmency H, Picard JC, Wang T. Fitness costs linked to dinitroaniline resistance mutation in Setaria. Heredity. 2011;107:80–6.
De Prado R, Dominguez C, Tena M. Characterization of triazine-resistant biotypes of common lambsquarters (Chenopodium album), hairy fleabane (Conyza bonaeriensis), and yellow foxtail (Setaria glauca) found in Spain. Weed Sci. 1989;37:1–4.
De Prado R, Diaz MA, Tena M. Differential tolerance to atrazine in four Setaria species. In: Proceedings of the EWRS symposium on integrated weed management in cereals, Helsinki, Finland; 4–6 Jun 1990. p. 61–8.
De Prado R, Lopez-Martinez N, Gonzalez-Gutierrez J. Identification of two mechanisms of atrazine resistance in Setaria faberi and Setaria viridis biotypes. Pest Biochem Physiol. 2000;67:114–24.
De Prado R, Osuna MD, Fischer AJ. Resistance to ACCase inhibitor herbicides in a green foxtail (Setaria viridis) biotype in Europe. Weed Sci. 2004;52:506–12.
De Wet JMJ, Oestry-Stidd LL, Cubero JI. Origins and evolution of foxtail millets (Setaria italica). J Agri Tradit Bot Appl. 1979;26:53–64.
Dekker J. The evolutionary biology of the foxtail (Setaria) species-group. In: Inderjit, editor. Weed biology and management. Dordrecht: Kluwer Academic; 2004. p. 65–113.
Délye C. Weed resistance to acetyl coenzyme A carboxylase inhibitors: an update. Weed Sci. 2005; 53:728–46.
Délye C, Wang T, Darmency H. An isoleucine-leucine substitution in chloroplastic acetyl-CoA carboxylase from green foxtail (Setaria viridis L Beauv) is responsible for resistance to the cyclohexanedione herbicide sethoxydim. Planta. 2002;214:421–7.
Délye C, Menchari Y, Michel S, Darmency H. Molecular bases for sensitivity to tubulin-binding herbicides in green foxtail. Plant Physiol. 2004;136:3920–32.
Délye C, Menchari Y, Michel S. A single polymerase chain reaction-based assay for simultaneous detection of two mutation conferring resistance to tubulin-binding herbicides in Setaria viridis. Weed Res. 2005;45:228–35.
Délye C, Jasienuk M, Le Corre V. Deciphering the evolution of herbicide resistance in weeds. Trends Genet. 2013;29:649–58.
Dong Z, Zhao H, He J, Huai J, Lin H, Zheng J, et al. Overexpression of a foxtail millet acetyl-CoA carboxylase gene in maize increases sethoxydim resistance and oil content. Afr J Biotechnol. 2011;10:3986–95.
Duke SO. Why have no new herbicide modes of action appeared in recent years? Pest Manag Sci. 2012;68:505–12.
Gasquez J, Compoint JC. Observation de chloroplastes résistants aux triazines chez une panicoidée, Setaria viridis L. Agronomie. 1981;1:923–6.
Gimenez-Espinosa R, Romera E, Tena M, De Prado R. Fate of atrazine in treated and pristine accessions of three Setaria species. Pest Biochem Physiol. 1996;56:196–207.
Hao HB, Cui HY, Li MZ. Germplasm innovation of herbicide resistant and colored summer millet. J Hebei Agric Sci. 2013;17:1–2.

Heap I. The international survey of herbicide resistant weeds. http://www.weedscience.org (2015). Accessed 12 Feb 2015.
Heap I, Morrison I. Resistance to aryloxyphenoxypropionate and cyclohexanedione herbicides in green foxtail (Setaria viridis). Weed Sci. 1996;44:25–30.
Jasieniuk M, Brûlé-Babel A, Morrison I. Inheritance of trifluralin resistance in green foxtail (Setaria viridis). Weed Sci. 1994;42:123–7.
Ji G, Du R, Hou S, Ceng R, Wang X, Zhao X. Genetics, development and application of cytoplasmic herbicide resistance in foxtail millet. Sci Agric Sin. 2006;39:879–85.
Jia X, Yuan J, Shi Y, Song Y, Wang G, Wang T, et al. A Ser-Gly substitution in plastid-encoded photosystem II D1 protein is responsible for atrazine resistance in foxtail millet (Setaria italica). Plant Growth Regul. 2007;52:81–9.
Konstantinovic B. Determination of triazine resistant biotypes of Setaria viridis. In: Proceedings of the BCPC conference: weeds, Brighton, UK; 12–15 Nov 2001. p. 607–12.
Laplante J, Rajcan I, Tardif FJ. Multiple allelic forms of acetohydroxyacid synthase are responsible for herbicide resistance in Setaria viridis. Theor Appl Genet. 2009;119:577–85.
Li MZ, Liu GR, Song K, Cui HY, Hao HB, Wang XM. Development status and analysis of hybrid millet seed industry. J Hebei Agric Sci. 2014;18:1–3.
Liu JJ, Zhao ZH. Cultivation technology and extension experiences of hybrid millet in Ethiopia. J Hebei Agric Sci. 2012;16:9–12.
Marles MAS, Devine MD, Hall JC. Herbicide resistance in Setaria viridis conferred by a less sensitive form of acetyl coenzyme A carboxylase. Pest Biochem Physiol. 1993;46:7–14.
Morrison IN, Todd BG, Nawolsky KM. Confirmation of trifluralin-resistant green foxtail (Setaria viridis) in Manitoba. Weed Technol. 1989;3:544–51.
Naciri Y, Darmency H, Belliard J, Dessaint F, Pernès J. Breeding strategy in foxtail millet, Setaria italica (LP Beauv), following interspecific hybridization. Euphytica. 1992;60:97–103.
Niu Y, Li Y, Shi Y, Song Y, Ma Z, Wang T, et al. AFLP mapping for the gene conferring sethoxydim resistance in foxtail millet (Setaria italica L Beauv). Acta Agric Sin. 2002;28:359–62.
Oliver LR, Schreiber MM. Differential selectivity of herbicides on six Setaria taxa. Weed Sci. 1971;19:428–31.
Reboud X, Till-Bottraud I. The cost of herbicide resistance measured by a competition experiment. Theor Appl Genet. 1991;82:690–6.
Ricroch A, Mousseau M, Darmency H, Pernès J. Comparison of triazine-resistant and -susceptible cultivated Setaria italica L (PB): growth and photosynthetic capacity. Plant Physiol Biochem. 1987;25:29–34.
Ritter RL, Kaufman LM, Monaco TJ, Novitzky WP, Moreland DE. Characterization of triazine-resistant giant foxtail (Setaria faberi) and its control in no-tillage corn (Zea mays). Weed Sci. 1989;37:591–5.
Santelmann PW, Meade JA. Variation in morphological characteristics and dalapon susceptibility within the species Setaria lutescens and S. faberii. Weeds. 1961;9:406–10.
Shanxi Academy of Agricultural Sciences. The cultivation of foxtail millet in China. In: Gu S, editor. Beijing: Agriculture Press; 1987.
Shi Y, Wang T, Li Y, Darmency H. Impact of transgene inheritance on the mitigation of gene flow between crops and their wild relatives: the example of foxtail millet. Genetics. 2008;180:969–75.
Shukla A, Leach GE, Devine MD. High-level resistance to sethoxydim conferred by an alteration in the target enzyme, acetyl-CoA carboxylase, in Setaria faberi and Setaria viridis. Plant Physiol Biochem. 1997;35:803–7.
Smeda RJ, Vaughn KC, Morrison IN. A novel pattern of herbicide cross-resistance in a trifluralin-resistant biotype of green foxtail [Setaria viridis (L.) Beauv.]. Pest Biochem Physiol. 1992;42:227–41.
Song XZ, Wang YW, Li HX, Tian G, Li ZH, Liu X. Breeding and the key cultivation techniques in herbicide-resistant millet Jingu 56. J Shanxi Agric Sci. 2014;42:115–8.
Stepenson GR, Dykstra MD, McLaren RD, Hamill AS. Agronomic practices influencing triazine-resistant weed distribution in Ontario. Weed Technol. 1990;4:199–207.
Stoltenberg DE, Wiederholt RJ. Giant foxtail (Setaria faberi) resistance to aryloxyphenoxypropionate and cyclohexanedione herbicides. Weed Sci. 1995;43:527–35.

Thompson Jr L. Metabolism of chloro s-triazine herbicides by Panicum and Setaria. Weed Sci. 1972;20:584–7.
Tian X, Darmency H. Rapid bidirectional allele-specific PCR identification for triazine resistance in higher plants. Pest Manag Sci. 2006;62:531–6.
Tian X, Délye C, Darmency H. Molecular evidence of biased inheritance of trifluralin herbicide resistance in foxtail millet. Plant Breed. 2006;125:254–8.
Tian BH, Wang JG, Li YJ, Zhang LX. Study on selection of suitable herbicide for hybrid millet. J Hebei Agric Sci. 2010;14:46–7.
Till-Bottraud I, Reboud X, Brabant P, Lefranc M, Rherissi B, Vedel F, et al. Outcrossing and hybridization in wild and cultivated foxtail mil-lets: consequences for the release of transgenic crops. Theor Appl Genet. 1992;83:940–6.
Vila-Aiub MM, Neve P, Powles SB. Fitness cost associated with evolved herbicide resistance alleles in plants. New Phytol. 2009;184:751–67.
Volenberg D, Stoltenberg D. Altered acetyl-coenzyme A carboxylase confers resistance to clethodim, fluazifop and sethoxydim in Setaria faberi and Digitaria sanguinalis. Weed Res. 2002a;42:342–50.
Volenberg DS, Stoltenberg DE. Giant foxtail (Setaria faberi) outcrossing and inheritance of resistance to acetyl-coenzyme A carboxylase inhibitors. Weed Sci. 2002b;50:622–7.
Volenberg DS, Stoltenberg DE, Boerboom CM. Biochemical mechanism and inheritance of cross-resistance to acetolactate synthase inhibitors in giant foxtail. Weed Sci. 2001;49:635–41.
Volenberg DS, Stoltenberg DE, Boerboom CM. Green foxtail (Setaria viridis) resistance to acetolactate synthase inhibitors. Phytoprotection. 2002;83:99–109.
Wang T, Darmency H. Dinitroaniline herbicide cross-resistance in resistant Setaria italica lines selected from interspecific cross with S. viridis. Pest Sci. 1997a;49:277–83.
Wang T, Darmency H. Inheritance of sethoxydim resistance in foxtail millet, Setaria italica (L). Beauv. Euphytica. 1997b;94:69–73.
Wang T, Darmency H. Cross-resistance to aryloxyphenoxypropionate and cyclohexanedione in foxtail millet (Setaria italica). Pest Biochem Physiol. 1998;59:81–8.
Wang RL, Dekker J. Weedy adaptation in Setaria spp. II. Variation in herbicide resistance in Setaria spp. Pest Biochem Physiol. 1995;51:99–116.
Wang T, Du R, Chen H, Darmency H, Fleury A. A new way of using herbicide resistant gene on hybrid utilization in foxtail millet. Sci Agric Sin. 1996a;29:96.
Wang T, Fleury A, Ma J, Darmency H. Genetic control of dinitroaniline resistance in foxtail millet (Setaria italica). J Hered. 1996b;87:423–6.
Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD. Construction of RFLP-based maps of foxtail millet, Setaria italica (L.) P. Beauv. Theor Appl Genet. 1998;96:31–36.
Wang T, Xin Z, Shi Y, Darmency H. A creation, evaluation and utilization of the new crop germplasm: herbicide resistance foxtail millet (Setaria italica). J China Agric Sci Technol. 2000;2:62–6.
Wang T, Zhao Z, Yan H, Li Y, Song Y, Ma Z, et al. Gene flow from cultivated herbicide-resistant foxtail millet to its wild relatives: a basis for risk assessment of the release of transgenic millet. Acta Agric Sin. 2001;27:681–7.
Wang T, Shi Y, Li Y, Song Y, Darmency H. Population growth rate of Setaria viridis in the absence of herbicide: resulting yield loss in foxtail millet Setaria italica. Weed Res. 2010a;50:228–34.
Wang T, Picard JC, Tian X, Darmency H. A herbicide-resistant ACCase 1781 Setaria mutant shows higher fitness than wild type. Heredity. 2010b;105:394–400.
Wang YW, Li HX, Tian G, Shi QX. Study on innovation and application of highly-male-sterile line with high outcrossing rate in millet. Sci Agric Sin. 2010c;43:680–9.
Zhou HZ, Hou SL, Song YF, Zhao Y, Dong L, Jia HY, et al. Impacts of monocotyledon weeds on millet yield loss in foxtail millet field. Chinese Agric Sci Bull. 2013;29:179–84.

# Chapter 16
# Genetic Determinants of Drought Stress Tolerance in *Setaria*

**Mehanathan Muthamilarasan and Manoj Prasad**

**Abstract** Cultivated foxtail millet (*Setaria italica*) and its wild progenitor (*S. viridis*) have collectively been considered as tractable model species for studying $C_4$ photosynthesis, stress biology, and biofuel traits. Being cultivated in arid and semi-arid tropics of the world, these species are well adapted to harsh environments such as drought, heat, and salinity. This adaptation or acclimation potential of *Setaria* spp. has drawn research interest, and attempts have been made to dissect the molecular mechanisms of stress tolerance. Compared to other stresses, drought response has been studied extensively in *S. italica* and many drought-responsive genes encoding for transcription factors, signaling molecules, and enzymes have been identified and characterized. Several genome-wide studies have reported on identification of stress-responsive gene family members, and speculated on the potential for expansion and neofunctionalization of paralogs in these gene families. In this context, this chapter discusses the key genetic determinants identified for stress tolerance in *S. italica* and demonstrates their use in improving drought tolerance. In addition, strategies for identification of genes underlying stress tolerance are also described. Little effort has so far been made towards understanding the stress-tolerance characteristics of *Setaria* as compared to studies reported in other crops. Comprehensive functional studies along with the use of integrated -omics approaches are required to elucidate the genetics and genomics of stress tolerance in *Setaria,* as it is important to develop climate change resilient crops to meet the growing demand for food and feed.

**Keywords** Foxtail millet (*Setaria italica*) • Green foxtail (*Setaria viridis*) • Drought stress • Gene expression • Transcription factors • Small RNAs • Next-generation sequencing

M. Muthamilarasan • M. Prasad, Ph.D. (✉)
Plant Molecular Genetics and Genomics, National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, JNU Campus, New Delhi, Delhi 110 067, India
e-mail: manoj_prasad@nipgr.ac.in



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_16

## 16.1 Introduction

Anthropogenic emissions of greenhouse gases have significantly contributed to global warming, resulting in increased atmospheric temperatures and unpredictable rainfall, which has serious impacts on agricultural productivity (IPCC 2014). Drought is one immediate outcome of global climate change and poses severe threats to agriculture, with the degree of its effect depending on onset time, duration, and intensity. Global temperature has been increased by 1.2 °C over the past century, and it is projected to rise by an average of 3 °C by 2010 (IPCC 2014), which would markedly affect the survival and yield of food crops. Occurrence of drought stress at the reproductive stage of field crops causes an average yield loss of more than 50 % (Venuprasad et al. 2007). In Australia, wheat production was halved after a ±2 °C temperature variation (Asseng et al. 2011). The adverse effects of climate change and decrease in arable land as well as a growing world population that is expected to reach nine billion by 2050 demands immediate action for doubling crop yields to meet the challenge of food and nutrition security (Karp and Richter 2011). Among cultivated crops, $C_3$ staple cereals such as wheat and rice are the worst affected by stresses imposed by climate change, particularly drought (Lal 2010). However, the productivity of underutilized grasses such as millets are affected less by drought as they are $C_4$ crops with better water use efficiency and are tolerant to a broad spectrum of biotic and abiotic stresses (Sadras et al. 2011). Furthermore, millets are cultivated in the arid and semiarid tropics of the world, where there is limited availability of rainfall and irrigation.

Most millets belong to the subfamily Panicoideae of the Poaceae and generally have a short life cycle, produce characteristic small grains, and can withstand dry and elevated temperature conditions. Most importantly, millets can survive on nutritionally poor soils with little compromise on yield. Among millets, *Setaria italica* (foxtail millet) and *S. viridis* (green foxtail), members of the tribe Paniceae, are considered as a model for studying $C_4$ photosynthesis and stress biology (Doust et al. 2009; Brutnell et al. 2010, 2015; Li and Brutnell 2011; Wang et al. 2011; Lata et al. 2013; Muthamilarasan and Prasad 2015). Reports have suggested the origin of cultivated *S. italica* from wild *S. viridis* ~11,000 years ago in Northern China (Yang et al. 2012). Presently, *S. italica* is being widely cultivated in tropical and subtropical regions of China, India, sub-Saharan Africa, and America for food and feed (Dwivedi et al. 2011). The prominent attributes of both *S. italica* and *S. viridis* (collectively, Setaria) include small diploid genomes (~490 Mb), short growing cycle (~90 days), small stature, $C_4$ traits, potential stress tolerance, and significant genetic colinearity with biofuel grasses and major cereals (Doust et al. 2009; Brutnell et al. 2010, 2015; Li and Brutnell 2011; Lata et al. 2013; Diao et al. 2014; Muthamilarasan and Prasad 2015). Therefore, considering Setaria as a model, the genomes of *S. italica* "Yugu1" and *S. viridis* "A10" were sequenced by the Joint Genome Institute—U.S. Department of Energy (Bennetzen et al. 2012) and the genome of *S. italica* "Zhang gu" and "A2" was sequenced by the Beijing Genomics Institute, China (Zhang et al. 2012; Lata and Prasad 2013a). In addition, the transcriptome of *S. italica* tissues such as root, leaf, stem, and spica (tassel) from young seedlings has also been sequenced (Zhang et al. 2012).

The ability of these plants to tolerate or avoid stress by acclimatization and adaptation suggests that a repository of genetic diversity essential for enhancing yield stability exists in the germplasm of both *S. italica* and its wild progenitor *S. viridis*. Therefore, identification of key genetic determinants of drought tolerance in Setaria using QTL mapping, association mapping, and screening by recurrent selection is imperative, as this would enable the transfer of genes into other crops using genomics-assisted breeding. The first comparative transcriptome of *S. italica* cultivar "Mar51" (drought tolerant; Zhang et al. 2005) in response to drought stress using subtracted cDNA libraries reported the up-regulation of 95 and 57 ESTs in roots and shoots, respectively (Zhang et al. 2007). These expressed sequence tags (ESTs) showed tissue-specific expression patterns, and it was deduced that activation of glycolysis metabolism in roots is the first response to drought stress (Zhang et al. 2007). Similar subtractive hybridization analyses were performed by Lata et al. (2011) and Puranik et al. (2011a) in stress-tolerant *S. italica* cv. "Prasad" during drought and salinity stress. These studies reported 327 and 159 differentially expressed transcripts in drought and salt stressed libraries, respectively. Comparative analysis of these differentially expressed transcripts from both libraries revealed that only 10 % of them are similar (Puranik et al. 2011a). This demonstrated the existence of gene sets which are distinct for drought and salt stress, suggesting the presence of unique tolerance mechanisms to circumvent each stress.

The release of the *S. italica* genome sequence has expedited investigations on stress-related studies, and many reports are now available on the identification of stress-responsive genes that might confer durable tolerance (Lata et al. 2011, 2014; Mishra et al. 2012a, b, 2013; Puranik et al. 2013; Muthamilarasan et al. 2014a, b; Wang et al. 2014a, b; Zhu et al. 2014; Yadav et al. 2015a; Kumar et al. 2015). Further characterization of these genes through transgene-based approaches to understand their role in molecular, cellular, and physiological processes of drought tolerance would enable the transfer of this knowledge to other related crop species. Though studies on deciphering the mechanism of drought tolerance in *S. italica* commenced a decade ago, similar investigations in *S. viridis* have yet not been reported (Muthamilarasan and Prasad 2015). In this context, this chapter presents an overview of research efforts made towards identifying and characterizing the genetic determinants of drought tolerance in *S. italica* and the strategies to transfer these genes/QTLs into modern crop germplasm using genomic approaches.

## 16.2 Drought-Responsive Transcription Factors

The response of plants to drought stress is a complex process involving multiple dynamic responses at physiological, biochemical, and molecular levels. The stress signal perceived is communicated via sophisticated signal transduction networks to initiate the activity of stress-responsive transcription factors (TFs). Comparative transcriptome analysis of drought-tolerant *S. italica* cv. "Prasad" using a subtractive hybridization technique has identified two important stress-responsive transcription

factors belonging to the DREB (dehydration-responsive element-binding proteins) and NAC (NAM, ATAF, and CUC) families (Lata et al. 2011). Transcript profiling of *SiDREB2* and *SiNAC2* genes using qRT-PCR in two *S. italica* cultivars with contrasting tolerance to dehydration stress (tolerant cv. "Prasad," susceptible cv. "Lepakshi") revealed significant up-regulation of these genes in the tolerant cultivar (Lata et al. 2011), suggesting putative involvement of these genes in stress-responsive mechanisms.

DREB is a subfamily of AP2/ERF (APETALA2/ethylene-responsive element-binding factor) TFs and participates in regulation of stress-responsive gene expression through ABA-independent pathways. This subfamily comprises two main subgroups, DREB1 and DREB2, which are involved in responses to chilling and drought, respectively (Lata and Prasad 2011). Cloning and characterization of *SiDREB2* revealed that it is a nuclear localized 234 amino acid protein (25.7 kDa) encoded by 1119 bp cDNA (Lata et al. 2011). The SiDREB2 protein comprises a 58 amino acid AP2/ERF DNA-binding domain along with two functional amino acids, valine and glutamic acid, at the 14th and 19th residues, respectively. These two amino acids were identified in the DBA-binding domain and deduced to be important for binding with their respective *cis*-elements. Sequence alignment showed that the AP2/ERF DNA-binding domain of SiDREB2 was highly conserved among AP2/ERF TFs of other Poaceae members (Lata et al. 2011). Expression profiling of *SiDREB2* in four *S. italica* cultivars (drought tolerant cv. "Prasad" and "IC-403579"; susceptible cv. "Lepakshi" and "IC-480117") during different time-points of drought, salinity, and cold stress showed the up-regulation of this gene (up to 12-fold) in tolerant cultivars in response to drought and salinity. The strong responsiveness of *SiDREB2* to drought and salinity in "Prasad" and "IC-403579" may be positively correlated to the tolerance behavior of these cultivars (Lata et al. 2011).

Sequence analysis of the *SiDREB2* gene in 43 contrasting *S. italica* cultivars identified a synonymous SNP associated with dehydration tolerance at the 558th base pair (an A/G transition) (Lata et al. 2011). An allele-specific marker (ASM) was developed from this SNP and validated in a core set of 170 *S. italica* accessions (Lata et al. 2011). The regression of lipid peroxidation (LP) and relative water content (RWC) on this ASM demonstrated that the *SiDREB2*-associated trait contributes to ~27 % and ~20 % of the total variation in LP and RWC, respectively (Lata and Prasad 2012, 2013b).

A genome-wide survey was conducted using in silico approaches based on the role of DREB TFs in stress response (Lata et al. 2014). The study revealed 171 AP2/ERF-encoding genes in the *S. italica* genome (Fig. 16.1), of which 48 were DREB TFs identified by phylogenetic and domain architecture analysis. Transcript profiling of candidate genes was performed in drought-tolerant foxtail millet cultivar "IC-403579" exposed to 20 % polyethylene glycol (PEG 6000) and 250 mM sodium chloride, with transcript abundance analyzed at 1 h (early) and 24 h (late) post-stress treatments. The DREB gene *SiAP2/ERF-002* was highly expressed in both phases of drought stress, and *SiAP2/ERF-084* as well as *SiAP2/ERF-090* was up-regulated in the late phase of drought and salinity stress. Hormonal treatment studies reported higher expression of *SiAP2/ERF-084* and *SiAP2/ERF-090* during the early phase of ethephone (converted

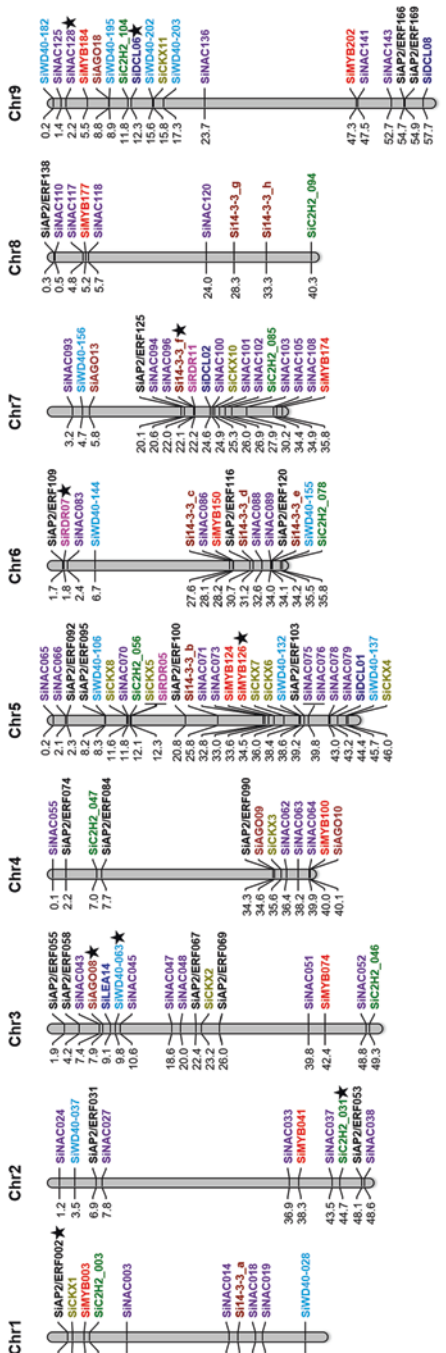


**Fig. 16.1** Chromosomal location of *S. italica* stress-responsive gene family members validated in response to drought stress. *Vertical bars* represent chromosomes and numbers in the *left* indicate the physical position (in Mbp). The respective gene names are provided in the *right* and the *color codes* correspond to each gene family. Genes which could be potential candidates for further functional characterization for drought stress tolerance are indicated by a star symbol (Refer Table 16.1)

into ethylene by the plant) and salicylic acid treatments, respectively. *SiAP2/ERF-002* was up-regulated in both early and late phases of ethephone and salicylic acid treatments (Lata et al. 2014). The study identified *SiAP2/ERF-002* as a potential candidate gene for further functional validation and overexpression studies with a view towards its utilization in crop improvement programs for stress tolerance.

NAC TFs are well known for their regulatory role in biotic as well as abiotic stress in many crop plants (Puranik et al. 2012). A subtractive hybridization study in *S. italica* also identified a significant up-regulation of *SiNAC2* in both drought (Lata et al. 2011) and salinity stress libraries (Puranik et al. 2011a). Molecular characterization of *SiNAC2* showed that the full length cDNA is 2051 bp with an open reading frame of 1386 bp encoding a protein of 462 amino acids (51.12 kDa). The full length SiNAC protein has a conserved NAC domain at its N-terminal (156 amino acids) along with a hypervariable C-terminal region. Using an electrophoretic mobility shift assay (EMSA), the DNA-binding site in the SiNAC2 protein has also been identified (Puranik et al. 2011b). Subcellular localization studies suggested that the SiNAC2 protein is membrane localized, and nuclear localization is also observed after deletion of the C-terminus. Expression profiling of *SiNAC* in *S. italica* cv. "Prasad" (drought tolerant) and "Lepakshi" (drought susceptible) in response to salinity and drought stress showed relatively higher levels of *SiNAC* transcripts in the tolerant cultivar, suggesting a positive role of SiNAC in stress response (Puranik et al. 2011b).

Similar reports are also available in other crop plants such as rice (*SNAC1*, Hu et al. 2006; *OsNAC045*, Zheng et al. 2009), soybean (*GmNAC2*, *GmNAC3*, *GmNAC4*, Pinheiro et al. 2009), and wheat (*TaNAC4*, Xia et al. 2010; *TaNAC2a*, Tang et al. 2012), substantiating the role of NAC TFs in diverse stress responses including drought tolerance. In view of this, a genome-wide analysis for identification and characterization of NAC TFs in *S. italica* was performed by Puranik et al. (2013). The study identified 147 *SiNAC* genes (Fig. 16.1), classified into 11 subfamilies. Of the 147 *SiNAC* genes, 50 candidate genes were chosen for quantitative expression analysis in response to various abiotic stresses. During drought stress, *SiNAC062, SiNAC064, SiNAC070,* and *SiNAC128* were observed to be up-regulated during the early phase, whereas *SiNAC024, SiNAC093*, *SiNAC100, SiNAC101,* and *SiNAC128* were up-regulated during the late phase of stress. The study identified *SiNAC128* as a potential candidate for further in-depth characterization (Puranik et al. 2013).

Availability of the *S. italica* draft genome sequence in the public domain has facilitated the identification and characterization of a few important stress-responsive TFs namely, MYB and $C_2H_2$-type zinc fingers. MYB and $C_2H_2$ proteins represent the largest TF families in plants, playing crucial roles in various developmental and stress-responsive processes (Ambawat et al. 2013). Considering their significance, comprehensive genome-wide surveys were conducted, which led to the identification of 209 and 124 gene family members of MYB and $C_2H_2$, respectively (Fig. 16.1). Phylogenetic analysis categorized SiMYB proteins into ten groups (I–X) and $SiC_2H_2$ proteins into five groups (I–V). Comparative analysis of SiMYB and $SiC_2H_2$ protein sequences with their orthologs in sorghum, maize, and rice showed a remarkable conservation in overall protein structure (Muthamilarasan et al. 2014a, b).

Expression analysis of *SiMYB* and *SiC$_2$H$_2$* candidate genes in response to abiotic stresses and hormone treatments using qRT-PCR revealed specific and/or overlapping expression patterns of these genes. In the case of *SiMYB* genes, 11 candidate genes were chosen for expression profiling and three genes (*SiMYB124*, *SiMYB126*, and *SiMYB150*) showed significant up-regulation during drought stress (Muthamilarasan et al. 2014a). Among the nine *SiC$_2$H$_2$* genes selected for expression studies, *SiC2H2_031* showed a gradual up-regulation with maximum expression at 48th hour (h) of drought stress, whereas *SiC2H2_78, SiC2H2_85*, and *SiC2H2_94* showed higher expression during the early phase of drought stress. Altogether, these studies have identified potential candidate genes for further functional validation and utilization in crop improvement programs for stress tolerance (Table 16.1).

In view of the role of TFs in modulating stress-responsive gene regulatory networks, TF-encoding genes in *S. italica* genome have been identified and the TFs in silico characterized (Bonthala et al. 2014). The study identified 2297 putative TFs and categorized them in 55 families. This information is available in the Foxtail millet Transcription Factor Database (http://59.163.192.91/FmTFDb/) in which complete details of the TFs are compiled, including their sequences, physical positions, tissue-specific gene expression data, gene ontologies, and phylogeny (Bonthala et al. 2014). This database will be useful in pinpointing candidate TFs for stress-related studies and for performing large-scale investigations.

Though TFs are reported to be effectual in enhancing stress tolerance of transgenic plants by regulating the expression of broad-spectrum stress-related genes, the lack of an efficient transformation system for expressing/overexpressing the candidate TFs is a bottleneck in *Setaria* genomics (Diao et al. 2014) (but see Chaps. 20 and 21). As a consequence, the detailed molecular, cellular, and physiological mechanisms responsible for variation in drought tolerance among foxtail millet lines have not yet been elucidated.

## 16.3 Stress-Responsive Proteins in Drought Tolerance

Other than transcription factors, various stress-responsive proteins have been reported to play roles in conferring tolerance to drought stress (Hasanuzzaman et al. 2013). One such protein is WD40, which largely functions as a platform for protein–protein interactions and is involved in several biological process, such as signal transduction, transcriptional regulation, protein modification, cytoskeleton assembly, vesicular trafficking, DNA damage and repair, cell death, and cell cycle progression (Mishra et al. 2012a; Zhang and Zhang 2015). Reports have shown the association of these proteins with abiotic stress tolerance in crop plants (Zhu et al. 2008; Lee et al. 2010; Mishra et al. 2012a; Kong et al. 2015). In *S. italica*, ESTs encoding for putative WD-domain containing proteins and 14-3-3 like proteins were identified from a salinity and dehydration stress subtractive cDNA library (Lata et al. 2011; Puranik et al. 2011a). The full length cDNA of *SiWD40* was deduced to be

**Table 16.1** List of stress-responsive gene families identified and characterized in response to drought stress in *Setaria italica*

| S. No. | Gene type | Name of gene family | Notation | No. of genes deduced in foxtail millet genome | No. of genes validated in response to drought stress | Potential candidate(s) for functional characterization (gene ID) | References |
|---|---|---|---|---|---|---|---|
| 1 | Transcription factors | NAM, ATAF, and CUC | NAC | 147 | 50 | *SiNAC128* (Si036695m) | Puranik et al. (2013) |
| 2 | | APETALA2/ethylene-responsive element-binding factor | AP2/ERF | 171 | 21 | *SiAP2/ERF-002* (Si018222m) | Lata et al. (2014) |
| 3 | | Myeloblastosis family | MYB | 209 | 11 | *SiMYB126* (Si005548m) | Muthamilarasan et al. (2014a, b) |
| 4 | | $C_2H_2$-type zinc finger | $C_2H_2$ ZF | 124 | 9 | *SiC2H2_031* (Si032711m) | Muthamilarasan et al. (2014a, b) |
| 5 | | WRKY domain-containing protein | WRKY | 110 | 12 | *SiWRKY068* (Si013326m) | Muthamilarasan et al. (2015a) |
| 6 | RNA silencing components | Dicer-like | DCL | 8 | 4 | *SiDCL06* (Si033906m) | Yadav et al. (2015a) |
| 7 | | Argonaute | AGO | 19 | 5 | *SiAGO08* (Si021147m) | Yadav et al. (2015a) |
| 8 | | RNA-dependent RNA polymerase | RDR | 11 | 3 | *SiRDR07* (Si015317m) | Yadav et al. (2015a) |
| 9 | Hormone metabolism | Cytokinin oxidase/dehydrogenase | CKX | 11 | 11 | – | Wang et al. (2014a, b) |
| 10 | Aldehyde metabolism | Aldehyde dehydrogenase | ALDH | 20 | 20 | *SiALDH2C4* (Si040073) | Zhu et al. (2014) |
| 11 | Multi-process | WD40 repeat proteins | WD40 | 225 | 13 | *SiWD40-063* (Si023326m) | Mishra et al. (2013) |
| 12 | | 14-3-3 proteins | 14-3-3 | 8 | 8 | *Si14-3-3_f* (Si010865m) | Kumar et al. (2015) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 13 | Secondary cell wall biosynthesis | Cellulose synthase | CesA | 14 | 2 | *SiCesA5* (Si028762m) | Muthamilarasan et al. (2015b) |
| | | Cellulose synthase-like | Csl | 39 | 0 | – | |
| | | Glucan synthase-like | Gsl | 12 | 2 | *SiGsl2* (Si016067m) | |
| | | Phenylalanine ammonia lyase | PAL | 10 | 2 | *SiPAL2* (Si016467m) | |
| | | Trans-cinnamate 4-hydroxylase | C4H | 3 | 1 | *SiC4H2* (Si022114m) | |
| | | 4-Coumarate CoA ligase | 4CL | 20 | 1 | *Si4CL10* (Si026881m) | |
| | | Hydroxycinnamoyl CoA:shikimate/quinate hydroxycinnamoyl transferase | HCT | 2 | 1 | *SiHCT1* (Si016926m) | |
| | | p-Coumaroyl shikimate 3′-hydroxylase | C3H | 2 | 0 | – | |
| | | Caffeoyl CoA 3-O-methyltransferase | CCoAOMT | 6 | 1 | *SiCCoAOMT3* (Si014344m) | |
| | | Ferulate 5-hydroxylase | F5H | 2 | 1 | *SiF5H2* (Si035174m) | |
| | | Caffeic acid O-methyltransferase | COMT | 4 | 1 | *SiCOMT2* (Si014900m) | |
| | | Cinnamoyl CoA reductase | CCR | 33 | 2 | *SiCCR7* (Si030374m) | |
| | | Cinnamyl alcohol dehydrogenase | CAD | 13 | 2 | *SiCAD6* (Si030413m) | |

– Data not available

1795 bp long with an ORF of 1314 bp, which encodes for a 437 amino acid protein (43.9 kDa). Protein modeling and analysis of SiWD40 revealed an eight blade β-propeller architecture at the C-terminus, with each blade comprising a four-stranded antiparallel β sheet (Mishra et al. 2012b). Transcript profiling of the *SiWD40* gene in *S. italica* stress tolerant cv. "IC-403579" at different time-points of drought stress revealed a steady-state transcript accumulation from early to late phase with maximum expression at the 48th h of drought stress. Subcellular localization studies have shown the localization of the SiWD40 protein in the nucleus, and EMSA along with transactivation assays have revealed the regulation of *SiWD40* gene expression by dehydration-responsive elements (DRE) (Mishra et al. 2012b).

Global analyses of WD40 protein-encoding genes in the *S. italica* genome showed the presence of 225 SiWD40 genes, classified into five subfamilies (Mishra et al. 2013) (Fig. 16.1). Expression analysis of 13 candidate WD40 genes in response to drought, salinity, and cold stresses has been performed. Among the candidate *SiWD40* genes, *SiWD40-028, SiWD40-037, SiWD40-063, SiWD40-106, SiWD40-156*, and *SiWD40-203* showed a gradual rise in expression levels and an average higher expression at 12–24 h (Mishra et al. 2013). This study suggests that SiWD40 proteins might play a prominent role in dynamically integrating multiple regulatory pathways mediating tolerance to abiotic stresses.

14-3-3 proteins are reported to regulate plant growth and development, and stress responses through protein–protein interactions, by binding with phosphoserine/phosphothreonine residues in the target proteins (Li and Dhaubhadel 2011; Denison et al. 2011). Bioinformatic prediction of 14-3-3 gene family members revealed the presence of 8 genes in the *S. italica* genome (Kumar et al. 2015) (Fig. 16.1). Further characterization revealed large variation in their structure, chromosomal localization, and protein properties, and in silico expression profiling indicated their higher expression in all the four investigated tissues of *S. italica* namely, roots, stems, leaves, and spikes. Comparative mapping to identify the orthologous genes in other grasses showed a high degree of conservation throughout the family. Subcellular localization studies showed differential localization of Si14-3-3_a, Si14-3-3_d, Si14-3-3_f, and Si14-3-3_h proteins within the cell. Si14-3-3_f was localized in cytoplasm and nuclear membrane, whereas the other three members were ubiquitously distributed (Kumar et al. 2015).

Transcript profiling of *Si14-3-3* genes in response to drought and salinity stress as well as ABA, SA, and MeJA treatments indicated that these genes have varied expression patterns. During drought stress, a relatively high expression of *Si14-3-3_a, Si14-3-3_c, Si14-3-3_d, Si14-3-3_f*, and *Si14-3-3_g* was reported at the early phase. Further downstream characterization indicated the interaction of Si14-3-3 with a nucleocytoplasmic shuttling phosphoprotein (SiRSZ21A) in a phosphorylation-dependent manner, demonstrating that Si14-3-3 might regulate the splicing events by binding with phosphorylated SiRSZ21A (Kumar et al. 2015). The demonstration of an interaction between Si14-3-3 and SiRSZ21A provides novel clues on the involvement of 14-3-3 proteins in splicing events. In this context, it would be interesting to investigate the protein–protein interaction behavior of 14-3-3 proteins during environmental stresses.

Cytokinins are reported to participate in various aspects of plant growth, development, and stress adaptations (Havlova et al. 2008), and the level of cytokinins is fine-tuned by cytokinin oxidase/dehydrogenases (CKXs) (Gajdosová et al. 2011). In maize and soybean, CKX genes have responded to drought and salinity stress (Vyroubalova et al. 2009; Le et al. 2012). Considering this, Wang et al. (2011) conducted a comprehensive genome-wide survey and identified 11 *SiCKX* genes in the *S. italica* genome. Phylogenetic analysis of SiCKX proteins with rice and *Arabidopsis* orthologs classified them in two groups. The relative transcript levels of *SiCKX* genes in germinating embryos under drought stress showed higher expression of all *SiCKX* genes except *SiCKX2* and *SiCKX11* (Wang et al. 2011), substantiating the role of *SiCKX* genes in response to drought stress.

Aldehyde dehydrogenases (ALDHs) are a conserved gene family encoding NAD $(P)^+$-dependent enzymes, which catalyze the irreversible oxidation of broad-spectrum endogenous and exogenous aromatic and aliphatic aldehydes into corresponding carboxylic acids (Yoshida et al. 1998). Studies have reported the involvement of these ALDHs in guarding plants from various biotic and abiotic stresses by indirectly detoxifying cellular reactive oxygen species (ROS) and/or reducing lipid peroxidation (Singh et al. 2013). In view of this, Zhu et al. (2014) performed a genome-wide analysis and identified 20 ALDH genes in *S. italica.* The study categorized these *SiALDH* genes into ten gene families and examined their duplication and divergence, chromosomal distribution, gene structure, and orthologous relationships with rice. Further, the *SiALDH* genes were subjected to quantitative expression analysis in response to drought, salt, high and low temperature, and hydrogen peroxide stress treatments. In the event of drought stress, all the *SiALDH2* genes were up-regulated except *SiALDH2C1, SiALDH3H2*, and *SiALDH11A1.* Of note, *SiALDH2C4* showed maximum expression at the 6th hour of drought stress (Zhu et al. 2014) and is a potential candidate for modifying drought stress response.

Late embryogenesis abundant (LEA) proteins are accumulated in seeds during the later stage of development before the desiccation phase, and these proteins are also reported to function in protecting the plants from environmental stresses (He et al. 2012; Liu et al. 2013). Reports have shown the accumulation of LEA proteins during drought stress, and their overexpression confers stable tolerance to water deficit in transgenic plants (Xu et al. 1996; Colmenero-Flores et al. 1999; Goyal et al. 2005; Tolleter et al. 2010). In *S. italica*, a novel member of the atypical subgroup 5C LEA gene named *SiLEA14* was functionally characterized (Wang et al. 2014a, b). The full length sequence of *SiLEA14* has 821 bp, encoding a 170 amino acid LEA protein (18.77 kDa), which is cytosol localized. Expression levels of *SiLEA14* under drought stress showed an immediate induction within 0.5 h of stress initiation and maximum expression was reported at 12 h. Overexpression of *SiLEA14* in *Arabidopsis* and *S. italica* enhanced the tolerance of transgenic plants to drought stress, with a higher accumulation of proline and sugar (Fig. 16.1; Wang et al. 2014a, b). The findings suggested that overexpression of *SiLEA14* gene in crop plants might improve tolerance to drought stress, but further experimental validation is required to support this conclusion.

Studies have been performed to identify the members of the RNA silencing machinery, such as Dicer-like (DCL), Argonaute (AGO), and RNA-dependent RNA polymerase (RDR) genes, and understand how they regulate gene expression during abiotic stress in *Arabidopsis* (Henderson et al. 2006), rice (Kapoor et al. 2008), maize (Qian et al. 2011), tomato (Bai et al. 2012), and poplar (Zhao et al. 2015). Similar efforts have identified 8 *DCL*, 19 *AGO*, and 11 *RDR* genes in *S. italica* (Yadav et al. 2015a) (Fig. 16.1). These genes have been characterized using in silico approaches, and expression profiling was performed for candidate genes (4 *DCL*, 5 *AGO*, and 3 *RDR*) in response to salinity and drought stress in two *S. italica* cultivars (cv. "IC04," stress tolerant; cv. "IC41," stress susceptible). The results found a differential expression pattern of candidate genes at two different time-points, stresses and cultivars, thus suggesting the participation of these genes in a complex molecular network of stress response (Yadav et al. 2015a). Significantly higher expression levels of *SiDCL01, SiDCL06, SiAGO08, SiAGO018*, and *SiRDR07* during drought stress suggests that these genes should be further functionally analyzed. Altogether, the identification, characterization, and expression profiling of stress-responsive protein-encoding genes in *S. italica* has established the putative role of respective proteins in complex networks of pathways to perform diverse physiological, molecular, and cellular functions in response to drought and other stresses. Though the expression profiling studies provide a clue to the role of these proteins in imparting stress tolerance (Table 16.1), comprehensive functional characterization is required to confirm their functionality and durability in conferring tolerance.

## 16.4 Small RNAs in Drought-Regulated Gene Expression

Small RNAs (sRNAs) are a part of noncoding RNAs, and they comprise two major classes namely, microRNAs (miRNAs) and endogenous small interfering RNAs (siRNAs). Both miRNAs and siRNAs are identified as modulators of gene expression at the post-transcriptional level and have emerged as key players in stress responses (Sunkar et al. 2007). miRNAs regulate the expression of the target transcript by binding to reverse complementary sequences, causing cleavage of the target RNA, whereas siRNAs bind to the target sequence in a similar manner and direct DNA methylation (Khraiwesh et al. 2012). Involvement of miRNAs in various biotic and abiotic stresses including drought (Zhao et al. 2007; Liu et al. 2008; Zhou et al. 2010), cold (Zhou et al. 2008), salinity (Liu et al. 2008; Sunkar et al. 2008), bacterial infection (Navarro et al. 2006), UV-B radiation (Zhou et al. 2007), and mechanical stress (Lu et al. 2005) have been well documented. Advances in high-throughput sequencing and small RNA profiling have facilitated the sequencing of small RNA libraries of drought stress samples for identification of drought-related miRNAs (Ding et al. 2013; Rajwanshi et al. 2014).

The genomic and CDS sequences of *S. italica* were analyzed for plant miRNA sequences, and 355 mature miRNAs (Sit-miR) were identified and classified into 53 families (Khan et al. 2014). Secondary structures and putative targets of Sit-miRs

were then identified, followed by chromosomal localization, comparative mapping, and tissue-specific expression profiling. Northern blot analysis and stem-loop RT-qPCR of candidate Sit-miRs in response to different abiotic stresses in two *S. italica* cultivars ("IC-403579," stress tolerant; "IC-480117," stress susceptible) were performed. The analysis showed up- and down-regulation of candidate Sit-miRs during various stresses. In the drought-tolerant cultivar, Sit-miR156c, Sit-miR397a, Sit-miR393, Sit-miR160d, and Sit-miR6248a were down-regulated. Up-regulation of Sit-miR162a, Sit-miR167b, and Sit-miR171b was also observed in the tolerant cultivar when compared to the expression levels of respective Sit-miRs in susceptible cultivars (Khan et al. 2014). The complete data of identified sit-miRs including chromosomal location, length, MFE, AMFE, sequences of pre-miRNA and mature miRNA, secondary structure, and target gene information have been made available to the global research community through an open access web resource, Foxtail millet miRNA Database (http://59.163.192.91/FmMiRNADb/index.html; Khan et al. 2014).

The use of two *S. italica* cultivars "IC-403579" (stress tolerant) and "IC-480117" (stress susceptible) in all the functional genomics studies discussed above encouraged us to construct four small RNA libraries from control and drought-stressed seedlings of these cultivars, which were then sequenced using the Illumina HiSeq 2000 platform (Yadav et al. 2016). A total of 55 known miRNAs (representing 23 miRNA families) and 136 novel miRNAs (representing 47 miRNA families) were identified in this study. Other downstream analyses such as chromosomal positioning, structure and target prediction, target annotation and validation, and expression profiling were performed, and a few candidate novel dehydration-responsive Sit-miRs were validated by stem-loop quantitative real-time PCR. Further functional characterization of these Sit-miRs are in progress (Yadav et al. 2016). Of note, this study showed differential expression pattern of Sit-miRs in response to drought, which may play an important role in providing the contrasting tolerance characteristics of these cultivars.

A similar bioinformatic approach was followed by Khan et al. (2014), with some modifications in the filtering criteria used by Han et al. (2014), and 271 miRNAs belonging to 44 families were predicted. The study identified 23 pairs of sense/antisense miRNAs and 18 miRNA clusters as well as 432 targets for 38 miRNA families. Of these, 43 miRNAs were chosen for tissue-specific expression profiling in *S. italica* leaves, roots, stems, and spikes, and five predicted targets of four miRNAs were experimentally validated using 5′-RLM-RACE (Han et al. 2014). In another study, two small RNA libraries constructed from shoot tissue of *S. italica* inbred line "Yugu1" were sequenced using the Illumina HighSeq 2000 platform, and 43 known miRNAs, 172 novel miRNAs, and 2 miRNA precursor candidates were identified (Yi et al. 2013). Targets of these miRNAs were predicted and annotated, followed by validation of candidate miRNAs by stem-loop RT-PCR in four tissues (Yi et al. 2013). Though these studies are insightful in understanding the miRNAome of *S. italica*, the role of identified miRNAs in response to drought and other stresses remains elusive.

Investigation of genome-wide transcriptome reconfiguration in *S. italica* challenged by drought stress was performed by Qi et al. (2013). RNA and sRNA libraries were constructed from drought stressed and unstressed (control) whole seedlings of

*S. italica* "Yugu1" and sequenced. Among the sRNAs, 24-nt (nucleotide) sRNAs were found to be predominant followed by 21, 22, and 23-nt sRNA. The study inferred that decreased levels of 24-nt siRNA around genic regions have a negative role in influencing gene expression in response to drought stress. Particularly, the differential expression analysis identified the maximum levels of 19 long noncoding RNAs during drought stress and, among these, two natural antisense transcripts (NATs of *Si003758m* and *Si038715m*) showed drought-regulated expression patterns (Qi et al. 2013). The generated raw reads from the studies of Yi et al. (2013) and Qi et al. (2013) are available in the NCBI SRA database under accession numbers SRA062640 and SRA062827.

Taken together, it is understood that identification and characterization of target genes is important for delineating the role of sRNAs. Prediction of sRNAs and their targets followed by their functional analysis will assist in understanding the complex miRNA- and siRNA-mediated regulatory networks controlling stress-responsive machinery in *Setaria*.

## 16.5 Strategies for Identifying Genetic Determinants of Drought Tolerance

Comprehensive molecular investigations have elucidated the role of a few known genes and gene families in stress responsiveness of *S. italica*, but so far no novel gene has been reported for stress tolerance. The genome annotation data of *S. italica* "Zhang gu" and "A2" has identified 1517 foxtail millet-specific gene families, of which 586 genes were annotated as "response to water" [GO:0009415, defined as any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus reflecting the presence, absence, or concentration of water] (Zhang et al. 2012). This demonstrates the existence of unexplored genetic determinants which could be responsible for stress-tolerance behavior and adaptation of Setaria to arid and semiarid environments. Furthermore, advances in next-generation sequencing (NGS) technologies and high-throughput analysis platforms provide an excellent opportunity to explore the genome and transcriptome of Setaria for pinpointing genes/alleles/QTLs responsible for drought tolerance.

Reports have indicated that drought tolerance is a complex trait dynamically controlled by numerous genes, and it is therefore imperative to identify the candidate genes to decipher the mechanism of drought response in Setaria, which would expedite genetic improvement either by molecular breeding or transgene-based approaches. Until now, genetic studies using approaches like subtractive cDNA hybridization, qRT-PCR, and cDNA microarray have been performed in Setaria (Diao et al. 2014; Muthamilarasan and Prasad 2015), but mapping of QTL at the gene level through map-based cloning has not been reported. In this context, NGS platforms can provide a comprehensive insight into sequence variations in the genome, and whole genome resequencing (WGR) could assist in detecting polymorphisms,

develop high-throughput markers, understand epigenetic modifications, identify splice variants, and perform expression profiling and DNA footprinting (Fig. 16.2) (Delseny et al. 2010). NGS is also used for identification of candidate genes and variants underlying important traits by linkage mapping, genome-wide association mapping, and genotyping-by-sequencing (Fig. 16.2) (Varshney et al. 2014).

Identifying and utilizing the sequence variation present in the genome is imperative for crop genetics and breeding. Availability of the *S. italica* genome sequence in public databases (Phytozome: http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sitalica; Foxtail millet Database: http://foxtailmillet.genomics.org.cn/page/species/index.jsp; PlantGDB: http://www.plantgdb.org/SiGDB/; *Setaria italica* Functional Genomics Database: http://structuralbiology.cau.edu.cn/SIFGD/) has enabled the development of high-throughput molecular markers such as microsatellites (Pandey et al. 2013; Kumari et al. 2013; Zhang et al. 2014), intron-length polymorphisms (Muthamilarasan et al. 2013), miRNA-based (Yadav et al. 2014) and transposable elements-based markers (Yadav et al. 2015b), and development of a marker database

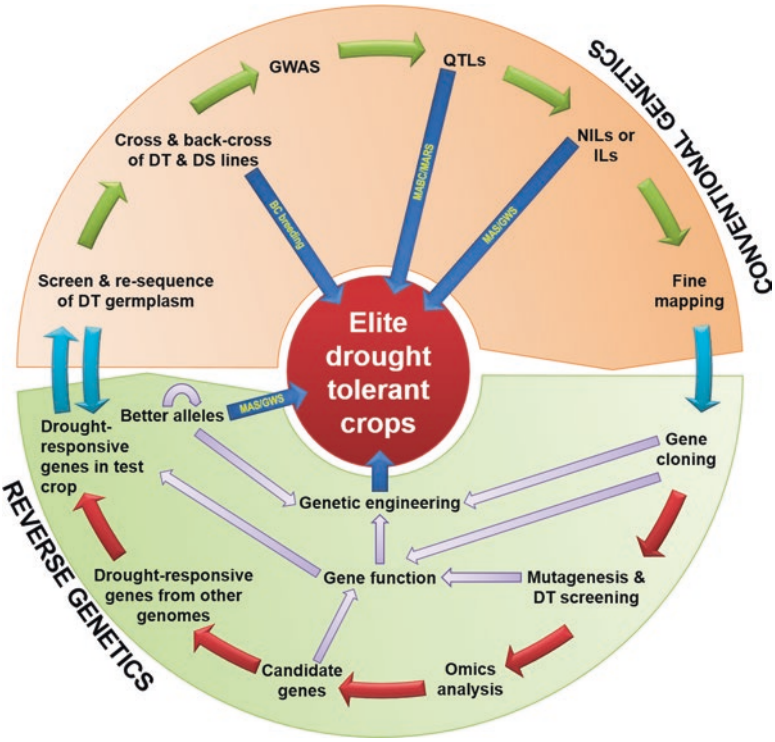


**Fig. 16.2** An integrated strategy of classical and reverse genetics for dissecting the drought-tolerance mechanisms and enhancing tolerance is illustrated. *DT* drought tolerance, *DS* drought susceptibility, *GWAS* genome-wide association study, *QTLs* quantitative trait loci, *NIL* near-isogenic line, *IL* isogenic line, *MABC* marker-assisted backcrossing, *MARS* marker-assisted recurrent selection, *MAS* marker-assisted selection, *GWS* genome-wide selection

(http://www.nipgr.res.in/foxtail.html; Suresh et al. 2013). The genome sequence information has also served as a reference in whole genome resequencing (WGR) of 916 *S. italica* accessions collected from different eco-geographical zones of the world and the construction of a high density haplotype map using 85 million single nucleotide polymorphisms, which revealed genomic variations among these accessions (Jia et al. 2013). Similarly, the *S. italica* genome sequence data facilitates WGR of cultivated and wild varieties of *Setaria* with contrasting phenotypes to identify novel genes/alleles/QTLs underlying drought response and to execute NGS-based genomics-assisted breeding for drought tolerance. Though the *S. viridis* genome has also been sequenced (Bennetzen et al. 2012), the lack of publically available sequence information has for a long period significantly impeded the development of genetic and genomic resources in this important model species. However, recent developments have made this data available in the web portal "Setariabase" (http://www.sviridis.org; Brutnell et al. 2015).

In addition to WGR, transcriptome sequencing could be useful for elucidating the transcriptional and post-transcriptional regulation of genes in response to drought stress and to understand the global expression pattern of the Setaria genome (Deyholos 2010). Transcriptomes of four *S. italica* tissues namely root, stem, leaf, and spike, as well as RNA-seq data of whole seedlings under drought conditions is already available in the NCBI SRA database (Zhang et al. 2012; Qi et al. 2013). In the case of *S. viridis,* RNA-seq data of leaf, stem, node, crown, root, spikelet, floret, and seed tissues at three developmental stages including seed germination, vegetative growth, and reproduction is available in NCBI SRA (Xu et al. 2013). Using these data as a reference, transcriptomes of *Setaria* accessions with distinctive phenotypes could be sequenced and compared to identify novel transcripts, which could be responsible for the trait-of-interest. Besides transcriptome sequencing, a few small RNA libraries have also been sequenced in control and stressed conditions, and miRNAs modulating drought stress-associated processes and gene networks identified (Yi et al. 2013; Qi et al. 2013; Yadav et al. 2016). Moreover, examining the DNA methylation profiles and small RNA profiles of drought-stressed libraries will facilitate the identification of genes and/or regions that are regulated by miRNA/siRNA-mediated DNA methylation, which could contribute to epigenetic inheritance of drought stress tolerance.

Proteomics offers versatile approaches for identifying drought-responsive proteins and corresponding genes (Kamal et al. 2010). More importantly, proteomics techniques are potent tools to delineate stress-responsive proteins and their corresponding genes even when genome sequence information is not available. A suggested pathway for such an analysis could include the resolution of total protein of *Setaria* cultivars subjected to drought stress and control conditions on two-dimensional gel electrophoresis, followed by in-gel digestion and MALDI-TOF mass spectrometry analysis. Subsequent in silico protein and nucleotide BLAST searches (e.g., using the MASCOT program; Brosch et al. 2009) would reveal differentially expressed proteins and their corresponding genes. Thus, proteomics could bridge the gap between transcriptome and metabolome and complement genomics approaches.

Recently, research has shifted from understanding the physiological and molecular responses of crop plants exposed to individual stress to those exposed to a combination of stresses, as numerous reports have demonstrated that the response of plants to concurrent stresses is unique and not directly extrapolated from the individual stress responses (Mittler 2006; Atkinson and Urwin 2012; Rasmussen et al. 2013; Suzuki et al. 2014; Ramegowda and Senthil-Kumar 2015). All the stress-related investigations conducted in Setaria were in response to different stresses applied individually and so far studies on the impact of combined/concurrent stress have not been performed. However, we can speculate that combinations of stresses could activate complex pathways controlled by different signaling events, which might be unique to Setaria. Therefore, research programs should focus on understanding the tolerance mechanism of Setaria to a combination of biotic and abiotic stress conditions, especially the stresses which mimic field conditions.

In the "-omics" era, identifying the genetic determinants for drought tolerance might not be a difficult task, but understanding the exact role of those genetic determinants in improving drought tolerance is challenging due to the lack of efficient Setaria transformation systems (Diao et al. 2014; Muthamilarasan and Prasad 2015; Brutnell et al. 2015). The availability of high-throughput genetic transformation systems has accelerated the maturation of rice and *Arabidopsis* genomics, but limited availability of such efficient protocols in *Setaria* has impeded further molecular studies in this model crop (Diao et al. 2014; Muthamilarasan and Prasad 2015). Further advances in transformation are reported in this volume (Chaps. 20 and 21). Irrespective of this, many candidate genes have been identified in *S. italica* which could be utilized in improving abiotic stress adaptation of other cereal crops.

## 16.6 Conclusions

Global climate change has caused irreversible damage to the earth's environment, which includes rise in atmospheric temperature, melting of glaciers, increase in sea levels, and changes in rainfall patterns. Emission of greenhouse gases due to extreme anthropogenic activities and extensive deforestation regimes accelerate the effects of climate change. These adverse conditions have severe impact on yields of crop plants, especially cereals (IPCC 2014) while, on the other hand, agricultural productivity needs to increase globally by an estimated 60 % by 2050 to meet the food and feed demands of a growing population (FAO 2015). Development of climate change resilient crops is the principal solution for this aggravating problem, but genes that enable growth and reproduction in adverse environments might have been lost in all the presently cultivated crops during domestication and improvement. Reports have shown that genetic diversity for stress tolerance, which enhances yield stability could be present in traditional landraces, wild relatives, and genetically close crop plants which are well adapted to adverse environments (Khoury et al. 2013; Atwell et al. 2014; Brutnell et al. 2015).

Being cereal crops, *Setaria* sp. and major cereals show high levels of synteny at the genome level, and, as discussed in the above sections, *S. italica* and *S. viridis* are tractable models for understanding stress biology and $C_4$ photosynthesis. Both species are tolerant to drought stress, which is the major impact of climate change, and in view of this, attempts have been made to identify the genetic determinants of drought stress using various molecular genetic and genomic approaches. At present, functional characterization is required to confirm the drought responsiveness of the identified genes and once confirmed, these could be introgressed into major cereals for generating elite cultivars with durable stress tolerance.

**Acknowledgements** Studies on millet genomics in Dr. Manoj Prasad's laboratory are supported by the Department of Biotechnology, Govt. of India and Core Grant of National Institute of Plant Genome Research (NIPGR), New Delhi, India. Mr. Mehanthan Muthamilarasan acknowledges the award of a Research Fellowship from the University Grants Commission, New Delhi, India. Authors also thank Ms. Jananee Jaishankar, NIPGR for critically reading this chapter.

## References

Ambawat S, Sharma P, Yadav NR, Yadav RC. MYB transcription factor genes as regulators for plant responses: an overview. Physiol Mol Biol Plants. 2013;19:307–21.

Asseng S, Foster I, Turner NC. The impact of temperature variability on wheat yields. Global Change Biol. 2011;17:997–1012.

Atkinson NJ, Urwin PE. The interaction of plant biotic and abiotic stresses: from genes to the field. J Exp Bot. 2012;63:3523–43.

Atwell BJ, Wang H, Scafaro AP. Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*? Plant Sci. 2014;215–216:48–58.

Bai M, Yang GS, Chen WT, Mao ZC, Kang HX, Chen GH, Yang YH, Xie BY. Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. Gene. 2012;501:52–62.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye CY, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM. Reference genome sequence of the model plant *Setaria*. Nat Biotechnol. 2012;30:555–61.

Bonthala VS, Muthamilarasan M, Roy R, Prasad M. FmTFDb: a foxtail millet transcription factors database for expediting functional genomics in millets. Mol Biol Rep. 2014;41:6343–8.

Brosch M, Yu L, Hubbard T, Choudhary J. Accurate and sensitive peptide identification with Mascot Percolator. J Proteome Res. 2009;8:3176–81.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu XG, Kellogg E, Van Eck J. *Setaria viridis*: a model for $C_4$ photosynthesis. Plant Cell. 2010;22:2537–44.

Brutnell TP, Bennetzen JL, Vogel JP. *Brachypodium distachyon* and *Setaria viridis*: model genetic systems for the grasses. Annu Rev Plant Biol. 2015;65:715–41.

Colmenero-Flores JM, Moreno LP, Smith CE, Covarrubias AA. Pvlea-18, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing regions of well-irrigated bean seedlings. Plant Physiol. 1999;120:93–104.

Delseny M, Han B, Hsing YI. High throughput DNA sequencing: the new sequencing revolution. Plant Sci. 2010;179:407–22.

Denison FC, Paul AL, Zupanska AK, Ferl RJ. 14-3-3 proteins in plant physiology. Semin Cell Dev Biol. 2011;22:720–7.
Deyholos MK. Making the most of drought and salinity transcriptomics. Plant Cell Environ. 2010;33:648–54.
Diao X, Schnable J, Bennetzen JL, Li J. Initiation of Setaria as a model plant. Front Agri Sci Eng. 2014;1:16–20.
Ding Y, Tao Y, Zhu C. Emerging roles of microRNAs in the mediation of drought stress response in plants. J Exp Bot. 2013;64:3077–86.
Doust AN, Kellogg EA, Devos KM, Bennetzen JL. Foxtail millet: a sequence-driven grass model system. Plant Physiol. 2009;149:137–41.
Dwivedi S, Upadhyaya H, Senthilvel S, Hash C, Fukunaga K, Diao X, Santra D, Baltensperger D, Prasad M. Millets: genetic and genomic resources. Plant Breed Rev. 2011;35:247–375.
FAO. The post-2015 development agenda and the millennium development goals. 2015. http://www.fao.org/post-2015-mdg/14-themes/climate-change/en/. Accessed 10 Apr 2015.
Gajdosová S, Spíchal L, Kamínek M, Hoyerová K, Novák O, Dobrev PI, Galuszka P, Klíma P, Gaudinová A, Zizková E, Hanus J, Dancák M, Trávnícek B, Pesek B, Krupicka M, Vanková R, Strnad M, Motyka V. Distribution, biological activities, metabolism, and the conceivable function of cis-zeatin-type cytokinins in plants. J Exp Bot. 2011;62:2827–40.
Goyal K, Walton LJ, Tunnacliffe A. LEA proteins prevent protein aggregation due to water stress. Biochem J. 2005;388:151–7.
Han J, Xie H, Sun Q, Wang J, Lu M, Wang W, Guo E, Pan J. Bioinformatic identification and experimental validation of miRNAs from foxtail millet (*Setaria italica*). Gene. 2014;546:367–77.
Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M. Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. Int J Mol Sci. 2013;14:9643–84.
Havlova M, Dobrev PI, Motyka V, Storchova H, Libus J, Dobra J, Malbeck J, Gaudinova A, Vankova R. The role of cytokinins in responses to water deficit in tobacco plants overexpressing trans-zeatin O-glucosyltransferase gene under 35S or SAG12 promoters. Plant Cell Environ. 2008;31:341–53.
He S, Tan L, Hu Z, Chen G, Wang G, Hu T. Molecular characterization and functional analysis by heterologous expression in E. coli under diverse abiotic stresses for OsLEA5, the atypical hydrophobic LEA protein from *Oryza sativa* L. Mol Genet Genomics. 2012;287:39–54.
Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE. Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat Genet. 2006;38:721–5.
Hu H, Dai M, Yao J, Xiao B, Xianghua L, Zhang Q, Xiong L. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. Proc Natl Acad Sci U S A. 2006;A103:12987–92.
IPCC. The Intergovernmental Panel on Climate Change: synthesis report. 2014. http://www.ipcc.ch/report/ar5/syr/. Accessed 15 Apr 2015.
Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, Chai Y, Yang L, Liu K, Lu H, Zhu C, Lu Y, Zhou C, Fan D, Weng Q, Guo Y, Huang T, Zhang L, Lu T, Feng Q, Hao H, Liu H, Lu P, Zhang N, Li Y, Guo E, Wang S, Wang S, Liu J, Zhang W, Chen G, Zhang B, Li W, Wang Y, Li H, Zhao B, Li J, Diao X, Han B. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (Setaria italica). Nat Genet. 2013;45:957-96.
Kamal AHM, Kim KH, Shin KH, Choi JS, Baik BK, Tsujimoto H, Heo HY, Park CS, Woo SH. Abiotic stress responsive proteins of wheat grain determined using proteomics technique. Aust J Crop Sci. 2010;4:196–208.
Kapoor M, Arora R, Lama T, Nijhawan A, Khurana JP, Tyagi AK, Kapoor S. Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analysis during reproductive development and stress in rice. BMC Genomics. 2008;9:451.
Karp A, Richter GM. Meeting the challenge of food and energy security. J Exp Bot. 2011;62:3263–71.

Khan Y, Yadav A, Suresh BV, Muthamilarasan M, Yadav CB, Prasad M. Comprehensive genome-wide identification and expression profiling of foxtail millet [*Setaria italica* (L.)] miRNAs in response to abiotic stress and development of miRNA database. Plant Cell Tiss Organ Cult. 2014;118:279–92.

Khoury CK, Greene S, Wiersema J, Maxted N, Jarvis A, Struik PC. An inventory of crop wild relatives of the United States. Crop Sci. 2013;53:1496–508.

Khraiwesh B, Zhu JK, Zhu J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim Biophys Acta. 2012;1819:137–48.

Kong D, Li M, Dong Z, Ji H, Li X. Identification of TaWD40D, a wheat WD40 repeat-containing protein that is associated with plant tolerance to abiotic stresses. Plant Cell Rep. 2015;34: 395–410.

Kumar K, Muthamilarasan M, Bonthala VS, Roy R, Prasad M. Unraveling 14-3-3 proteins in C4 panicoids with emphasis on model plant *Setaria italica* reveals phosphorylation-dependent subcellular localization of RS splicing factor. PLoS One. 2015;10:e0123236.

Kumari K, Muthamilarasan M, Misra G, Gupta S, Subramanian A, Parida SK, Chattopadhyay D, Prasad M. Development of eSSR-markers in *Setaria italica* and their applicability in studying genetic diversity, cross-transferability and comparative mapping in millet and non-millet species. PLoS One. 2013;8:e67742.

Lal R. Managing soils for a warming earth in a food-insecure and energy-starved world. J Plant Nutr Soil Sci. 2010;173:4–15.

Lata C, Prasad M. Role of DREBs in regulation of abiotic stress responses in plants. J Exp Bot. 2011;14:4731–48.

Lata C, Prasad M. Validation of an allele-specific marker associated with dehydration stress tolerance in a core set of foxtail millet accessions. Plant Breed. 2012;132:496–9.

Lata C, Prasad M. *Setaria* genome sequencing: an overview. J Plant Biochem Biotechnol. 2013a;22:257–60.

Lata C, Prasad M. Association of an allele-specific marker with dehydration stress tolerance in foxtail millet suggests *SiDREB2* to be an important QTL. J Plant Biochem Biotechnol. 2013b;23:119–22.

Lata C, Bhutty S, Bahadur RP, Majee M, Prasad M. Association of an SNP in a novel DREB2-like gene SiDREB2 with stress tolerance in foxtail millet [*Setaria italica* (L.)]. J Exp Bot. 2011;62:3387–401.

Lata C, Gupta S, Prasad M. Foxtail millet: a model crop for genetic and genomic studies in bioenergy grasses. Crit Rev Biotechnol. 2013;33:328–43.

Lata C, Mishra AK, Muthamilarasan M, Bonthala VS, Khan Y, Prasad M. Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in foxtail millet (*Setaria italica* L.). PLoS One. 2014;9:e113092.

Le DT, Nishiyama R, Watanabe Y, Vankova R, Tanaka M, Seki M, Ham LH, Yamaguchi-Shinozaki K, Shinozaki K, Tran LSP. Identification and expression analysis of cytokinin metabolic genes in soybean under normal and drought conditions in relation to cytokinin levels. PLoS One. 2012;7:e42411.

Lee S, Lee J, Paek K-H, Kwon S-Y, Cho H, Kim S, Park J. A novel WD40 protein, *BnSWD1*, is involved in salt stress in *Brassica napus*. Plant Biotechnol Rep. 2010;4:165–72.

Li P, Brutnell TP. *Setaria viridis* and *Setaria italica*, model genetic systems for the panicoid grasses. J Exp Bot. 2011;62:3031–7.

Li X, Dhaubhadel S. Soybean 14-3-3 gene family: identification and molecular characterization. Planta. 2011;233:569–82.

Liu HH, Tian X, Li YJ, Wu CA, Zheng CC. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. RNA. 2008;14:836–43.

Liu Y, Wang L, Xing X, Sun L, Pan J, Kong X, Zhang M, Li D. ZmLEA3, a multifunctional group 3 LEA protein from maize (*Zea mays* L.), is involved in biotic and abiotic stresses. Plant Cell Physiol. 2013;54(6):944–59.

Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL. Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell. 2005;17:2186–203.

Mishra AK, Puranik S, Bahadur RP, Prasad M. The DNA-binding activity of an AP2 protein is involved in transcriptional regulation of a stress-responsive gene, *SiWD40*, in foxtail millet. Genomics. 2012a;100:252–63.

Mishra AK, Puranik S, Prasad M. Structure and regulatory networks of WD40 protein in plants. J Plant Biochem Biotechnol. 2012b;21:32–9.

Mishra AK, Muthamilarasan M, Khan Y, Parida SK, Prasad M. Genome-wide investigation and expression analyses of WD40 protein family in the model plant foxtail millet (*Setaria italica* L.). PLoS One. 2013;9:e86852.

Mittler R. Abiotic stress, the field environment and stress combination. Trends Plant Sci. 2006;11: 15–9.

Muthamilarasan M, Prasad M. Advances in *Setaria* genomics for genetic improvement of cereals and bioenergy grasses. Theor Appl Genet. 2015;128:1–14.

Muthamilarasan M, Venkata Suresh B, Pandey G, Kumari K, Parida SK, Prasad M. Development of 5123 intron-length polymorphic markers for large-scale genotyping applications in foxtail millet. DNA Res. 2013;21:41–52.

Muthamilarasan M, Khandelwal R, Yadav CB, Bonthala VS, Khan Y, Prasad M. Identification and molecular characterization of MYB transcription factor superfamily in C4 model plant foxtail millet (*Setaria italica* L.). PLoS One. 2014a;9:e109920.

Muthamilarasan M, Bonthala VS, Mishra AK, Khandelwal R, Khan Y, Roy R, Prasad M. $C_2H_2$-type of zinc finger transcription factors in foxtail millet define response to abiotic stresses. Funct Integr Genomics. 2014b;14:531–43.

Muthamilarasan M, Bonthala VS, Khandelwal R, Jaishankar J, Shweta S, Nawaz K, Prasad M. Global analysis of WRKY transcription factor superfamily in Setaria identifies potential candidates involved in abiotic stress signaling. Front Plant Sci. 2015a;6:910.

Muthamilarasan M, Khan Y, Jaishankar J, Shweta S, Lata C, Prasad M. Integrative analysis and expression profiling of secondary cell wall genes in C4 biofuel model Setaria italica reveals targets for lignocellulose bioengineering. Front Plant Sci. 2015b;6:965.

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JD. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science. 2006;312: 436–9.

Pandey G, Misra G, Kumari K, Gupta S, Parida SK, Chattopadhyay D, Prasad M. Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. DNA Res. 2013;20:197–207.

Pinheiro GL, Marques CS, Costa MDBL, Reis PAB, Alves MS, Carvalho CM, Fietto LG, Fontes EPB. Complete inventory of soybean NAC transcription factors: sequence conservation and expression analysis uncover their distinct roles in stress response. Gene. 2009;444:10–23.

Puranik S, Jha S, Srivastava PS, Sreenivasulu N, Prasad M. Comparative transcriptome analysis of contrasting foxtail millet cultivars in response to short-term salinity stress. J Plant Physiol. 2011a;168:280–7.

Puranik S, Kumar K, Srivastava PS, Prasad M. Electrophoretic mobility shift assay reveals a novel recognition sequence for *Setaria italica* NAC protein. Plant Signal Behav. 2011b;6:1588–90.

Puranik S, Sahu PP, Srivastava PS, Prasad M. NAC proteins: regulation and role in stress tolerance. Trends Plant Sci. 2012;17:369–81.

Puranik S, Sahu PP, Mandal SN, Venkata Suresh B, Parida SK, Prasad M. Comprehensive genome-wide survey, genomic constitution and expression profiling of the NAC transcription factor family in foxtail millet (*Setaria italica* L.). PLoS One. 2013;8:e64594.

Qi X, Xie S, Liu Y, Yi F, Yu J. Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. Plant Mol Biol. 2013;83:459–73.

Qian Y, Cheng Y, Cheng X, Jiang H, Zhu S, Cheng B. Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. Plant Cell Rep. 2011;30:1347–63.

Rajwanshi R, Chakraborty S, Jayanandi K, Deb B, Lightfoot DA. Orthologous plant microRNAs: microregulators with great potential for improving stress tolerance in plants. Theor Appl Genet. 2014;127:2525–43.

Ramegowda V, Senthil-Kumar M. The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. J Plant Physiol. 2015;176:47–54.

Rasmussen S, Barah P, Suarez-Rodriguez MC, Bressendorff S, Friis P, Costantino P, Bones AM, Nielsen HB, Mundy J. Transcriptome responses to combinations of stresses in Arabidopsis. Plant Physiol. 2013;161:1783–94.

Sadras V, Grassini P, Steduto P. Status of water use efficiency of main crops. In: The state of world's land and water resources for food and agriculture (SOLAW). Rome: FAO, London: Earthscan. 2011. http://www.fao.org/fileadmin/templates/solaw/files/thematic_reports/TR_07_web.pdf. Accessed 2 Feb 2015.

Singh S, Brocker C, Koppaka V, Chen Y, Chen Y, Jackson BC, Matsumoto A, Thompson DC, Vasiliou V. Aldehyde dehydrogenases in cellular responses to oxidative/electrophilic stress. Free Radic Biol Med. 2013;56:89–101.

Sunkar R, Chinnusamy V, Zhu J, Zhu JK. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci. 2007;12:301–9.

Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu JK. Identification of novel and candidate miRNAs in rice by high throughput sequencing. BMC Plant Biol. 2008;8:25.

Suresh BV, Muthamilarasan M, Misra G, Prasad M. FmMDb: a versatile database of foxtail millet markers for millets and bioenergy grasses research. PLoS One. 2013;8:e71418.

Suzuki N, Rivero RM, Shulaev V, Blumwald E, Mittler R. Abiotic and biotic stress combinations. New Phytol. 2014;203:32–43.

Tang Y, Liu M, Gao S, Zhang Z, Zhao X, Zhao C, Zhang F, Chen X. Molecular characterization of novel TaNAC genes in wheat and overexpression of TaNAC2a confers drought tolerance in tobacco. Physiol Plant. 2012;144:210–24.

Tolleter D, Hincha DK, Macherel D. A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. Biochim Biophys Acta. 2010;1798:1926–33.

Varshney RK, Terauchi R, McCouch SR. Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. PLoS Biol. 2014;12:e1001883.

Venuprasad R, Lafitte HR, Atlin GN. Response to direct selection for grain yield under drought stress in rice. Crop Sci. 2007;47:285–93.

Vyroubalova S, Vaclavikova K, Tureckova V, Novak O, Smehilova M, Hluska T, Ohnoutkova L, Frebort I, Galuszka P. Characterization of new maize genes putatively involved in cytokinin metabolism and their expression during osmotic stress in relation to cytokinin levels. Plant Physiol. 2009;151:433–47.

Wang L, Peterson RB, Brutnell TP. Regulatory mechanisms underlying $C_4$ photosynthesis. New Phytol. 2011;190:9–20.

Wang M, Li P, Li C, Pan Y, Jiang X, Zhu D, Zhao Q, Yu J. SiLEA14, a novel atypical LEA protein, confers abiotic stress resistance in foxtail millet. BMC Plant Biol. 2014a;14:290.

Wang Y, Liu H, Xin Q. Genome-wide analysis and identification of cytokinin oxidase/dehydrogenase (CKX) gene family in foxtail millet (*Setaria italica*). Crop J. 2014b;2:244–54.

Xia N, Zhang G, Sun YF, Zhu L, Xu LS, Chen XM, Liu B, Yu YT, Wang XJ, Huang LL, Kang ZS. TaNAC8, a novel NAC transcription factor gene in wheat, responds to stripe rust pathogen infection and abiotic stresses. Physiol Mol Plant Pathol. 2010;74:394–402.

Xu D, Duan X, Wang B, Hong B, Ho T, Wu R. Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol. 1996;110:249–57.

Xu J, Li Y, Ma X, Ding J, Wang K, Wang S, Tian Y, Zhang H, Zhu XG. Whole transcriptome analysis using next-generation sequencing of model species *Setaria viridis* to support C4 photosynthesis research. Plant Mol Biol. 2013;83:77–87.

Yadav CB, Muthamilarasan M, Pandey G, Khan Y, Prasad M. Development of novel microRNA-based genetic markers in foxtail millet for genotyping applications in related grass species. Mol Breed. 2014;34:2219–24.

Yadav CB, Muthamilarasan M, Pandey G, Prasad M. Identification, characterization and expression profiling of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in foxtail millet. Plant Mol Biol Rep. 2015a;33:43–55.

Yadav CB, Bonthala VS, Muthamilarasan M, Pandey G, Khan Y, Prasad M. Genome-wide development of transposable elements-based markers in foxtail millet and construction of an integrated database. DNA Res. 2015b;22:79–90.

Yadav A, Khan Y, Prasad M. Dehydration-responsive miRNAs in foxtail millet: genome-wide identification, characterization and expression profiling. Planta. 2016;243:749–66.

Yang X, Wan Z, Perry L, Lu H, Wang Q, Zhao C, Li J, Xie F, Yu J, Cui T, Wang T, Li M, Ge Q. Early millet use in northern China. Proc Natl Acad Sci U S A. 2012;109:3726–30.

Yi F, Xie S, Liu Y, Qi X, Yu J. Genome-wide characterization of microRNA in foxtail millet (*Setaria italica*). BMC Plant Biol. 2013;13:212.

Yoshida A, Rzhetsky A, Hsu LC, Chang C. Human aldehyde dehydrogenase gene family. Eur J Biochem. 1998;251:549–57.

Zhang C, Zhang F. The multifunctions of WD40 proteins in genome integrity and cell cycle progression. J Genomics. 2015;3:40–50.

Zhang JP, Wang MY, Bai YF, Jia JP, Wang GY. Rapid evaluation on the drought tolerance of foxtail millet at seedling stage. J Plant Genet Resour. 2005;6:59–62.

Zhang J, Liu T, Fu J, Zhu Y, Jia J, Zheng J, Zhao Y, Zhang Y, Wang G. Construction and application of EST library from *Setaria italica* in response to dehydration stress. Genomics. 2007;90:121–31.

Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, Xie M, Zeng P, Yue Z, Wang W, Tao Y, Bian C, Han C, Xia Q, Peng X, Cao R, Yang X, Zhan D, Hu J, Zhang Y, Li H, Li H, Li N, Wang J, Wang C, Wang R, Guo T, Cai Y, Liu C, Xiang H, Shi Q, Huang P, Chen Q, Li Y, Wang J, Zhao Z, Wang J. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol. 2012;30:549–54.

Zhang S, Tang C, Zhao Q, Li J, Yang L, Qie L, Fan X, Li L, Zhang N, Zhao M, Liu X, Chai Y, Zhang X, Wang H, Li Y, Li W, Zhi H, Jia G, Diao X. Development of highly polymorphic simple sequence repeat markers using genome-wide microsatellite variant analysis in Foxtail millet [*Setaria italica* (L.) P. Beauv]. BMC Genomics. 2014;2:15.

Zhao B, Liang R, Ge L, Li W, Xiao H, Lin H, Ruan K, Jin Y. Identification of drought-induced microRNAs in rice. Biochem Biophys Res Commun. 2007;354:585–90.

Zhao K, Zhao H, Chen Z, Feng L, Ren J, Cai R, Xiang Y. The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: gene structure, gene expression, phylogenetic analysis and evolution. J Genet. 2015;94(2):317–21.

Zheng X, Chen B, Lu G, Han B. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. Biochem Biophys Res Commun. 2009;379:985–9.

Zhou X, Wang G, Zhang W. UV-B responsive microRNA genes in *Arabidopsis thaliana*. Mol Syst Biol. 2007;3:103.

Zhou X, Wang G, Sutoh K, Zhu JK, Zhang W. Identification of cold-inducible microRNAs in plants by transcriptome. Biochim Biophys Acta. 2008;1779:780–8.

Zhou L, Liu Y, Liu Z, Kong D, Duan M, Luo L. Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. J Exp Bot. 2010;61:4157–68.

Zhu J, Jeong JC, Zhu Y, Sokolchik I, Miyazaki S, Zhu JK, Hasegawa PM, Bohnert HJ, Shi H, Yun DJ, Bressan RA. Involvement of Arabidopsis HOS15 in histone deacetylation and cold tolerance. Proc Natl Acad Sci U S A. 2008;105:4945–50.

Zhu C, Ming C, Zhao-shi X, Lian-cheng L, Xue-ping C, You-zhi M. Characteristics and expression patterns of the aldehyde dehydrogenase (ALDH) gene superfamily of foxtail millet (*Setaria italica* L.). PLoS One. 2014;9:e101136.

# Chapter 17
# *Setaria viridis* as a Model for $C_4$ Photosynthesis

**Carla Coelho, Pu Huang, and Thomas P. Brutnell**

**Abstract** Climate change compounded with dwindling arable lands, and population growth has presented a grand challenge for plant science to develop higher yielding varieties grown on fewer acres with fewer inputs. The most productive and photosynthetically efficient crops are $C_4$ grasses that have evolved mechanisms to concentrate $CO_2$ and reduce photorespiration in hot and dry environments. This improved biochemistry is most often achieved through compartmentalization of photosynthetic activities into two cell types, the bundle sheath (BS) and mesophyll (M) cells. BS and M cells are arranged in files of concentric rings around the vasculature, in a pattern known as Kranz anatomy. Ambitious efforts to engineer $C_4$ traits into $C_3$ crops have the potential to increase the yield of rice by up to 50 %, improve nitrogen use efficiency, and reduce water loss (Hibberd et al., Curr Opin Plant Biol 11: 228–31, 2008). To accelerate these engineering efforts, a better understanding is needed of the developmental and regulatory mechanisms underlying $C_4$ photosynthesis. Here, we discuss the many advantages of using *Setaria viridis* as a model system for dissecting $C_4$ photosynthesis (Brutnell et al., Plant Cell 22:2537–44, 2010; Annu Rev Plant Biol 66:465–85, 2015), particularly as a platform for gene discovery and to unravel the complex genetic basis underlying the anatomical and biochemical innovations of this trait.

## 17.1 *Setaria viridis* to Fast Track Unanswered Questions Underlying $C_4$ Photosynthesis

The biochemistry of $C_4$ photosynthesis was first described 50 years ago by Hal Hatch and Roger Slack (Hatch and Slack 1966), who proposed that a four-carbon compound is produced in a series of reactions to reduce $CO_2$ into sugars. In subsequent years, elegant biochemistry coupled to anatomical characterizations of leaf tissues revealed that the pathway is partitioned into two specialized cells (Rathnam and Edwards 1975; Jenkins and Boag 1985; Hatch et al. 1988). In the vast majority of $C_4$ plants, the M cells

C. Coelho, Ph.D. (✉) • P. Huang • T.P. Brutnell
Donald Danforth Plant Science Center, 975 North Warson Road, St Louis, MO 63132, USA
e-mail: ccoelho@danforthcenter.org



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_17

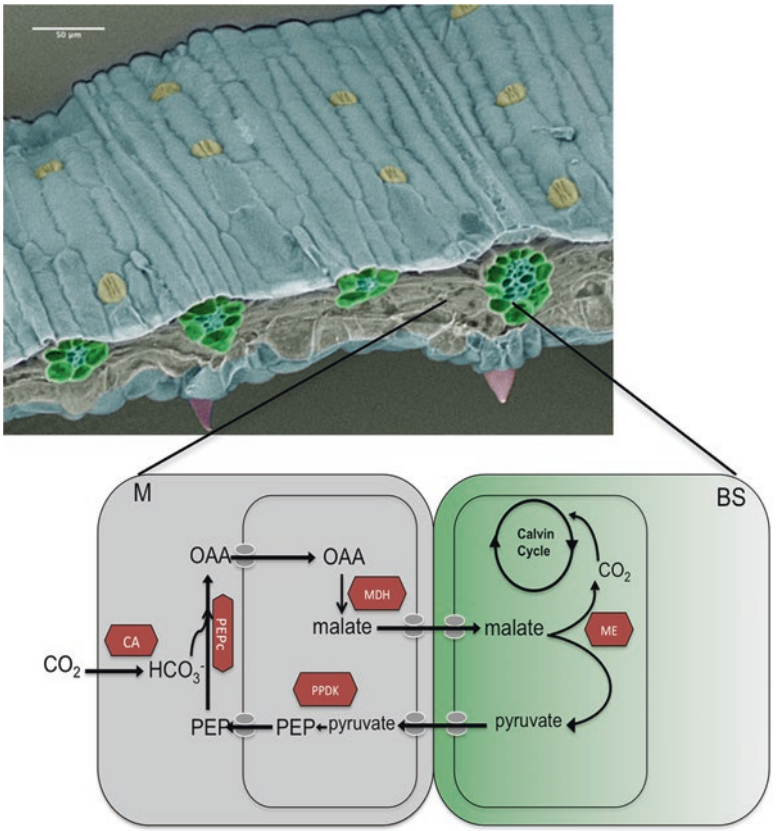


**Fig. 17.1** SEM of a *Setaria* leaf cross-section and a schematic view of the carbon flow of the $C_4$ NADP-ME subtype cycle between a M and BS cell: The NADP-ME subtype cycle starts when $CO_2$ enters the cytoplasm of the M cells and is converted into bicarbonate by the CA (carbonic anhydrase) enzyme and fixed by PEPCase, resulting in OAA (oxaloacetate). OAA moves into the M chloroplasts where it is converted to malate by the MDH (NADP-Malate dehydrogenase) enzyme. The four-carbon malate moves into the BS chloroplasts where it is decarboxylated by NADP-ME to release the $CO_2$ that is then refixed by RuBisCO in the Calvin Cycle. Pyruvate is generated after the decarboxylation of malate, shuttled back into the M chloroplasts, and is used to regenerate phosphoenolpyruvate (PEP) by the PPDK (pyruvate orthophosphate dikinase) enzyme. PEP is transported back into the M cytoplasm where it will be used as a substrate by PEPCase and restart the cycle. SEM image was pseudo-colored to highlight the different cell types. Vascular core: *blue*; Bundle sheath: *green*; Mesophyll: *gray*; Stomata: *yellow*; Trichomes: *pink*

are the sites of primary carbon fixation and use phosphoenolpyruvate carboxylase (PEPCase) to generate oxaloacetate (OAA) from bicarbonate ($HCO_3^-$) and phosphoenol pyruvate (PEP). OAA is further reduced to malate or transaminated to form aspartate which is then shuttled into the adjacent BS cells for decarboxylation and fixation by Rubisco (Fig. 17.1; Furbank 2011). This two-step process allows $C_4$ species to concentrate $CO_2$ in the vicinity of RuBisCO (Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase) thus minimizing photorespiration (Sage et al. 2012; Sage 2013). There are

three primary decarboxylating enzymes in the $C_4$ carbon cycle: NADP-ME (NADP-malic enzyme), NAD-ME (NAD-malic enzyme), and PCK (phosphoenol pyruvate carboxykinase), each of which accumulate preferentially in the BS. Interestingly, although all $C_4$ species utilize PEPCase for primary carbon fixation, one or more decarboxylases are utilized (Muhaidat et al. 2007). NADP-ME is the preferred decarboxylating enzyme in maize, sorghum, Miscanthus, sugarcane, and Setaria (Table 17.1).

Kranz anatomy has been postulated to be one of the first steps necessary for the establishment of $C_4$ photosynthesis (Sage et al. 2013). In $C_4$ grasses, the differentiation of the BS and M cells in leaf blades follows the development of the leaf vasculature (Langdale et al. 1988). Although rice and many $C_3$ grasses share a similar progression of vascular development, the $C_3$ leaf contains fewer small and intermediate vascular bundles, and a largely non-photosynthetic BS (Nelson and Langdale 1989) (Fig. 17.2). As photosynthetic cells differentiate in $C_4$ species, plastids acquire unique morphologies and functions in the BS and M cells, concomitant with the accumulation of differential enzymatic profiles and photosynthetic activities and metabolites (Majeran et al. 2008, 2010; Majeran and van Wijk 2009; Pick et al. 2011; Wang et al. 2014; Li et al. 2010). Importantly, the developing grass leaf is an excellent system to map the initiation and maturation of the photosynthetic machinery, as it follows a basipetal pattern of cellular differentiation with the oldest cells at the tip of the leaf differentiating prior to the youngest cells that are at the base of the leaf blade. This specialized aspect of grass leaf development enables the use of discrete transverse leaf sections to serve as a proxy for a developmental time series in order to interrogate changes that occur at the cellular and subcellular level during the establishment of $C_4$ photosynthesis.

In recent years, an expanded set of tools has been applied to better understand $C_4$ photosynthesis in panicoid grasses that include global proteomic (Majeran et al. 2008, 2010) and transcriptomic surveys (Wang et al. 2013, 2014; Li et al. 2010). By examining the shifting profiles of protein/transcript abundance across different tissue types and/or developmental stages, these studies have implicated hundreds of putative $C_4$ genes. These candidate genes likely encode enzymes, transporters, and transcriptional regulators necessary for the installation and maintenance of a complex $C_4$ photosynthetic machinery, but functional validation of these genes remains very limited (Hibberd and Covshoff 2010; Burgess and Hibberd 2015). In maize, for example, dissection of early stages of the foliar and husk leaf development enabled the identification of several transcription factors likely involved in the increased vein density found in $C_4$ tissues (Wang et al. 2013). However, experimental validations of $C_4$ candidate genes are rarely described (Studer et al. 2014; Furumoto et al. 2011; Bailey et al. 2000; Furbank et al. 1996; Von Caemmerer et al. 1997, 2005; Pengelly et al. 2012). Historically, this was partially a consequence of the genetic limitations of the model systems used where transformation technology or mutant screens were arduous and time consuming (Brutnell et al. 2015). One of the major challenges with maize as a model system is the long generation time, large genome size, and the need for field or greenhouse space to grow plants to maturity, in addition to the high cost and time frame required for plant transformation and regeneration. To date, the pri-

**Table 17.1** Characteristics of different $C_4$ grasses.

| Species name | Common name | Major decarboxylase | $C_4$ origin | M pyruvate transport | Genome size (1C) | Chromosome (2n) |
|---|---|---|---|---|---|---|
| *Zea mays* | Maize | NADP-ME (PCK) | Andropogon | Proton dependent | ~2.5 Gb | 20 |
| *Sorghum biocolor* | Sorghum | NADP-ME | Andropogon | Proton dependent | ~770 Mb | 20 |
| *Saccharum officinarum* | Sugarcane | NADP-ME | Andropogon | Proton dependent | ~930 Mb | 80, varies |
| *Setaria italica/viridis* | Foxtail millet/green foxtail | NADP-ME | Paniceae | Sodium dependent | ~500 Mb | 18 |
| *Panicum virgatum* | Switchgrass | NAD-ME | Paniceae | Sodium dependent | ~750 Gb | 36,72 |
| *Megathyrsus maximus* | Guinea grass | PCK | Paniceae | Sodium dependent | ~500 Mb | 18,36, varies |

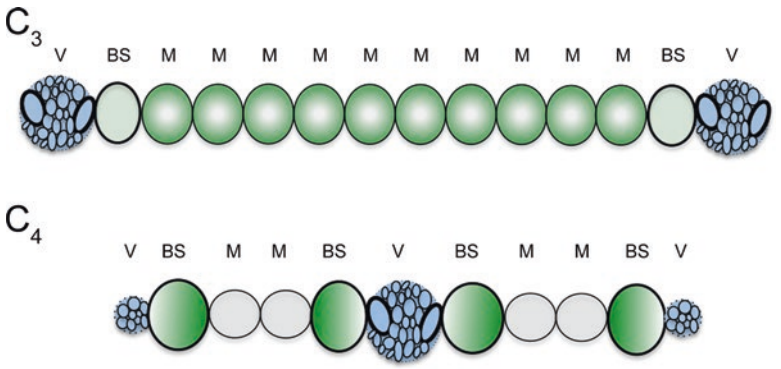


**Fig. 17.2** Schematic view of the leaf cell organization in a $C_3$ (A) and $C_4$ (B) grass. Extent of plastid differentiation and localization in the BS is indicated by intensity of green shading

mary obstacle in functional validation of $C_4$ regulators has been the lack of a simple and efficient transformation platform. For example, to introduce a single construct into an inbred line of maize may take as long as 1 year to produce seed from a few transgenic maize plants and cost $5000 per construct (http://agron-www.agron.iastate.edu/ptf/pricing/AgrobacteriumMaizeTransformation.aspx). Propagation of the plants in greenhouses adds additional costs. For instance, at the Donald Danforth Plant Science Center, space requirements and costs to propagate 100 maize transgenic plants are approximately ~350 $ft^2$ of greenhouse, with a monthly cost of ~$1250. In contrast, *Setaria* transformation can be performed in as little as 3 months, and the cost to propagate and maintain *Setaria* transgenics is considerably lower, as no more than three trays would be required for the same number of transgenics.

### *17.1.1 Exploiting Grass Synteny in Gene Discovery*

A surprisingly high proportion (87%) of transcripts share the same BS/M specificity between maize and *Setaria* (John et al. 2014) indicating a strong evolutionary conservation between these two $C_4$ lineages. This highly conserved profile of gene expression between *Setaria* and a distantly related panicoid grass suggests that similar conservation in cell-type enriched gene expression profiles will be shared with other panicoids such as sorghum and sugarcane. These findings also suggest that similar genetic regulatory modules have been co-opted for $C_4$ development and function. *Setaria* (a member of the Paniceae tribe) diverged from maize and other Andropogoneae grasses (e.g., Miscanthus, sugarcane, and sorghum) approximately 25 mya and does not share the same origin of $C_4$ with the Andropogoneae grasses. However, both *Setaria* and maize utilize NADP-ME as the major decarboxylase in the BS. Moreover, the same orthologous NADP-ME gene family member was recruited in Setaria and maize. Interestingly, switchgrass and guinea grass that acquired $C_4$ at approximately the same time as *Setaria*, use NAD-malic enzyme (NAD-ME) and phosphoenolpyruvate carboxykinase

(PCK) as their primary decarboxylases, respectively (Brutnell et al. 2010; Emms et al. 2016). From an engineering perspective, cross-species differences and similarities in the $C_4$ biochemistry observed in *Setaria* and other $C_4$ grasses could provide opportunities to discover pathways or genes in one grass and engineer these traits into another. The choice of decarboxylases can be quite versatile in some species (Furbank 2011); in maize for example, PCK accounts for a small but significant proportion (~25 %) of the $C_4$ decarboxylation activities (Wingler et al. 1999). In fact, when the primary pathway of malate import is impaired through the mutation of the malate transporter DCT2, maize lines can survive to maturity and produce seeds through the utilization of the PCK pathway (Weissmann et al. 2016). There also appears to be significant plasticity in the recruitment of transporters. One early study (Aoki et al. 1992) suggests that Andropogoneae species use a sodium-dependent pyruvate transport system, in contrast to other PACMAD species that use a proton-dependent pyruvate transport system. Expressing enzymes and transporters of the PCK pathway and the pyruvate transport system in *Setaria* would elucidate whether plants expressing multiple decarboxylases may benefit from the increased metabolite flow in the $C_4$ cycle and potentially enhance the overall photosynthetic rates.

Although much progress has been made in elucidating the biochemistry underlying $C_4$ photosynthesis, the regulatory networks that drive the formation of Kranz anatomy have remained much more elusive. To date, the growth hormones, gene networks, and regulatory circuits that underlie the differentiation of the BS and M cells remain undefined (Langdale 2011). As several cell-specific transcriptomic datasets have been published in the last few years (Li et al. 2010; Wang et al. 2013; John et al. 2014; Chang et al. 2012), it is now possible to develop testable hypotheses and validate them through functional genomics in *Setaria viridis*.

Perhaps one of the most intriguing hypotheses put forward in recent years (Slewinski et al. 2012, 2014; Slewinski 2013) is that SHORTROOT (SHR) and SCARECROW (SCR) homologs have been recruited to help pattern BS/M differentiation in $C_4$ grasses from an ancestral $C_3$ program that patterns both the BS in the shoot and the endodermis in the root (Nakajima et al. 2001; Cui et al. 2007; Di Laurenzio et al. 1996; Wysocka-Diller et al. 2000). Recent studies suggest that the development of BS cells in *Arabidopsis thaliana* (Cui et al. 2014) and in maize (Slewinski et al. 2012, 2014) require members of the SHR and SCR families. In the roots, the regulatory mechanisms by which SHR/SCR specify the endodermal cell layer in *Arabidopsis* is conserved among species (Wu et al. 2014), an indication that the function of these genes in BS/M differentiation might be also conserved among $C_4$ grasses as well. Regulatory networks often involve the function of a number of factors belonging to large gene families, for example, the SHR/SCR genes belong to the GRAS family (Rim et al. 2011), that in maize comprises more than 80 members. It has been shown that other members of the GRAS family, such as SCL23 and RGA1 may also function as cell specification factors in the leaves and roots (Cui et al. 2014; Moubayidin et al. 2016). Similarly, candidates belonging to the INDETERMINATE DOMAIN (IDD) family of transcription factors that are also known to modulate SHR/SCR activity in the roots (Long et al. 2015; Moreno-Risueno et al. 2015) were highlighted in independent comparative transcriptomics

studies in maize and *Setaria* as potential Kranz anatomy regulators (Wang et al. 2013; John et al. 2014). It is likely that *S. viridis* will become an important model system to dissect the potentially complex interactions among the IDD, SHR, and SCR family members.

## 17.2 Conclusions

Given the wealth of recent -omics datasets, the $C_4$ community is now poised to functionally dissect the regulatory and developmental controls of $C_4$ photosynthesis. *Setaria viridis* has emerged as an excellent genetic model plant for $C_4$ grasses with great potential to accelerate the pace of discovery in understanding $C_4$ biochemistry and Kranz anatomy (Martin et al. 2016; Fouracre et al. 2014). However, as the majority of $C_4$ -omics datasets have been generated in maize and sorghum (Majeran and van Wijk 2009; Majeran et al. 2010; Pick et al. 2011; Li et al. 2010; Wang et al. 2009, 2013; Ponnala et al. 2014), there remains a need to more deeply explore anatomical, proteomic, and gene expression profiles of *S. viridis*. Gene/anatomical/proteomic atlases of the *S. viridis* leaf will provide an important foundation for future *S. viridis* studies. As detailed in Chaps. 18 and 19, potential mutants in nearly all genes of *S. viridis* are likely segregating in NMU-mutagenized lines. The ability to perform crosses (Jiang et al. 2013) in *Setaria viridis* and prolific seed production are attractive features, especially when coupled to the genetic resources available with precise genome engineering through the use of CRISPR/Cas9 technology. This allows one to rapidly target candidate genes that may be involved in the biochemistry and developmental regulation of $C_4$ photosynthesis. Ultimately, fundamental questions will be more rapidly answered using *Setaria viridis*, and the discoveries translated to important agricultural crops, such as rice (Hibberd et al. 2008) and bioenergy feedstocks.

## References

Aoki N, Ohnishi J, Kanai R. Two different mechanisms for transport of pyruvate into mesophyll chloroplasts of $C_4$ plants—a comparative study. Plant Cell Physiol. 1992;33:805–9.

Bailey KJ, Battistelli A, Dever LV, Lea PJ, Leegood RC. Control of $C_4$ photosynthesis: effects of reduced activities of phosphoenolpyruvate carboxylase on $CO_2$ assimilation in *Amaranthus edulis* L. J Exp Bot. 2000;51:339–46.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu XG, et al. Setaria viridis: a model for C4 photosynthesis. Plant Cell. 2010;22:2537–44.

Brutnell TP, Bennetzen JL, Vogel JP. *Brachypodium distachyon* and *Setaria viridis* : model genetic systems for the grasses. Annu Rev Plant Biol. 2015;66:465–85. doi:10.1146/annurev-arplant-042811-105528.

Burgess SJ, Hibberd JM. Insights into $C_4$ metabolism from comparative deep sequencing. Curr Opin Plant Biol. 2015;25:138–44.

Chang Y-M, Liu W-Y, Shih AC-C, Shen M-N, Lu C-H, Lu M-YJ, et al. Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. Plant Physiol. 2012;160:165–77.

Cui H, Levesque MP, Vernoux T, Jung JW, Paquette AJ, Gallagher KL, et al. An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. Science. 2007;316:421–5.

Cui H, Kong D, Liu X, Hao Y. SCARECROW, SCR-LIKE 23 and SHORT-ROOT control bundle sheath cell fate and function in Arabidopsis thaliana. Plant J. 2014;78:319–27.

Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, et al. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. Cell. 1996;86:423–33.

Emms DM, Covshoff S, Hibberd JM, Kelly S. Independent and parallel evolution of new genes by gene duplication in two origins of C4 photosynthesis provides new insight into the mechanism of phloem loading in C4 species. Mol Biol Evol. 2016.

Fouracre JP, Ando S, Langdale JA. Cracking the Kranz enigma with systems biology. J Exp Bot. 2014;65:3327–39.

Furbank RT. Evolution of the $C_4$ photosynthetic mechanism: are there really three $C_4$ acid decarboxylation types? J Exp Bot. 2011;62:3103–8.

Furbank RT, Chitty JA, Von Caemmerer S, Jenkins C. Antisense RNA inhibition of RbcS gene expression reduces Rubisco level and photosynthesis in the C4 plant Flaveria bidentis. Plant Physiol. 1996;111:725–34.

Furumoto T, Yamaguchi T, Ohshima-Ichie Y, Nakamura M, Tsuchida-Iwata Y, Shimamura M, et al. A plastidial sodium-dependent pyruvate transporter. Nature. 2011;476:472–5.

Hatch MD, Slack CR. Photosynthesis by sugar-cane leaves. A new carboxylation reaction and the pathway of sugar formation. Biochem J. 1966;101:103–11.

Hatch MD, Agostino A, Burnell JN. Photosynthesis in phosphoenolpyruvate carboxykinase-type C4 plants: activity and role of mitochondria in bundle sheath cells. Arch Biochem Biophys. 1988;261:357–67.

Hibberd JM, Covshoff S. The regulation of gene expression required for $C_4$ photosynthesis. Annu Rev Plant Biol. 2010;61:181–207.

Hibberd JM, Sheehy JE, Langdale JA. Using C4 photosynthesis to increase the yield of rice-rationale and feasibility. Curr Opin Plant Biol. 2008;11:228–31.

Jenkins CL, Boag S. Isolation of bundle sheath cell chloroplasts from the NADP-ME type C(4) plant Zea mays: capacities for CO(2) assimilation and malate decarboxylation. Plant Physiol. 1985;79:84–9.

Jiang H, Barbier H, Brutnell T. Methods for performing crosses in Setaria viridis, a new model system for the grasses. J Vis Exp. 2013.

John CR, Smith-Unna RD, Woodfield H, Covshoff S, Hibberd JM. Evolutionary convergence of cell-specific gene expression in independent lineages of $C_4$ grasses. Plant Physiol. 2014;165:62–75.

Langdale JA. C4 cycles: past, present, and future research on C4 photosynthesis. Plant Cell. 2011;23:3879–92.

Langdale JA, Rothermel BA, Nelson T. Cellular pattern of photosynthetic gene expression in developing maize leaves. Genes Dev. 1988;2:106–15.

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, et al. The developmental dynamics of the maize leaf transcriptome. Nat Genet. 2010;42:1060–7.

Long Y, Smet W, Cruz-Ramirez A, Castelijns B, de Jonge W, Mahonen AP, et al. Arabidopsis BIRD zinc finger proteins jointly stabilize tissue boundaries by confining the cell fate regulator SHORT-ROOT and contributing to fate specification. Plant Cell. 2015;27:1185–99.

Majeran W, van Wijk KJ. Cell-type-specific differentiation of chloroplasts in C4 plants. Trends Plant Sci. 2009;14:100–9.

Majeran W, Zybailov B, Ytterberg AJ, Dunsmore J, Sun Q, van Wijk KJ. Consequences of $C_4$ differentiation for chloroplast membrane proteomes in maize mesophyll and bundle sheath cells. Mol Cell Proteomics. 2008;7:1609–38.

Majeran W, Friso G, Ponnala L, Connolly B, Huang M, Reidel E, et al. Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. Plant Cell. 2010;22:3509–42.

Martin AP, Palmer WM, Brown C, Abel C, Lunn JE, Furbank RT, et al. A developing Setaria viridis internode: an experimental system for the study of biomass generation in a C4 model species. Biotechnol Biofuels. 2016;9:45.

Moreno-Risueno MA, Sozzani R, Yardimci GG, Petricka JJ, Vernoux T, Blilou I, et al. Transcriptional control of tissue formation throughout root development. Science. 2015;350:426–30.

Moubayidin L, Salvi E, Giustini L, Terpstra I, Heidstra R, Costantino P, et al. A SCARECROW-based regulatory circuit controls Arabidopsis thaliana meristem size from the root endodermis. Planta. 2016;243(5):1159–68.

Muhaidat R, Sage RF, Dengler NG. Diversity of Kranz anatomy and biochemistry in C4 eudicots. Am J Bot. 2007;94:362–81.

Nakajima K, Sena G, Nawy T, Benfey PN. Intercellular movement of the putative transcription factor SHR in root patterning. Nature. 2001;413:307–11.

Nelson T, Langdale JA. Patterns of leaf development in C4 plants. Plant Cell. 1989;1:3–13.

Pengelly JJ, Tan J, Furbank RT, von Caemmerer S. Antisense reduction of NADP-malic enzyme in Flaveria bidentis reduces flow of CO2 through the C4 cycle. Plant Physiol. 2012;160:1070–80.

Pick TR, Brautigam A, Schluter U, Denton AK, Colmsee C, Scholz U, et al. Systems analysis of a maize leaf developmental gradient redefines the current C4 model and provides candidates for regulation. Plant Cell. 2011;23:4208–20.

Ponnala L, Wang Y, Sun Q, van Wijk KJ. Correlation of mRNA and protein abundance in the developing maize leaf. Plant J. 2014;78:424–40.

Rathnam CK, Edwards GE. Intracellular localization of certain photosynthetic enzymes in bundle sheath cells of plants possessing the C4 pathway of photosynthesis. Arch Biochem Biophys. 1975;171:214–25.

Rim Y, Huang L, Chu H, Han X, Cho WK, Jeon CO, et al. Analysis of Arabidopsis transcription factor families revealed extensive capacity for cell-to-cell movement as well as discrete trafficking patterns. Mol Cells. 2011;32:519–26.

Sage RF. Photorespiratory compensation: a driver for biological diversity. Plant Biol (Stuttg). 2013;15:624–38.

Sage RF, Sage TL, Kocacinar F. Photorespiration and the evolution of $C_4$ photosynthesis. Annu Rev Plant Biol. 2012;63:19–47.

Sage TL, Busch FA, Johnson DC, Friesen PC, Stinson CR, Stata M, et al. Initial events during the evolution of C4 photosynthesis in C3 species of Flaveria. Plant Physiol. 2013;163:1266–76.

Slewinski TL. Using evolution as a guide to engineer kranz-type c4 photosynthesis. Front Plant Sci. 2013;4:212.

Slewinski TL, Anderson AA, Zhang C, Turgeon R. Scarecrow plays a role in establishing Kranz anatomy in maize leaves. Plant Cell Physiol. 2012;53:2030–7.

Slewinski TL, Anderson AA, Price S, Withee JR, Gallagher K, Turgeon R. Short-root1 plays a role in the development of vascular tissue and kranz anatomy in maize leaves. Mol Plant. 2014; 7:1388–92.

Studer AJ, Gandin A, Kolbe AR, Wang L, Cousins AB, Brutnell TP. A limited role for carbonic anhydrase in C4 photosynthesis as revealed by a ca1ca2 double mutant in maize. Plant Physiol. 2014;165:608–17.

Von Caemmerer S, Millgate A, Farquhar GD, Furbank RT. Reduction of ribulose-1,5-bisphosphate carboxylase/oxygenase by antisense RNA in the C4 plant Flaveria bidentis leads to reduced assimilation rates and increased carbon isotope discrimination. Plant Physiol. 1997;113:469–77.

von Caemmerer S, Hendrickson L, Quinn V, Vella N, Millgate AG, Furbank RT. Reductions of Rubisco activase by antisense RNA in the C4 plant Flaveria bidentis reduces Rubisco carbamylation and leaf photosynthesis. Plant Physiol. 2005;137:747–55.

Wang X, Gowik U, Tang H, Bowers JE, Westhoff P, Paterson AH, et al. Comparative genomic analysis of $C_4$ photosynthetic pathway evolution in grasses. Genome Biol. 2009;10:R68.

Wang P, Kelly S, Fouracre JP, Langdale JA. Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of $C_4$ Kranz anatomy. Plant J. 2013;75:656–70.

Wang L, Czedik-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, et al. Comparative analyses of $C_4$ and $C_3$ photosynthesis in developing leaves of maize and rice. Nat Biotechnol. 2014; 32:1158–65.

Weissmann S, Ma F, Furuyama K, Gierse J, Berg H, Shao Y, et al. Interactions of C4 subtype metabolic activities and transport in maize are revealed through the characterization of DCT2 mutants. Plant Cell. 2016;28:466–84.

Wingler A, Walker RP, Chen Z-H, Leegood RC. Phosphoenolpyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. Plant Physiol. 1999;120:539–46.

Wu S, Lee CM, Hayashi T, Price S, Divol F, Henry S, et al. A plausible mechanism, based upon Short-Root movement, for regulating the number of cortex cell layers in roots. Proc Natl Acad Sci U S A. 2014;111:16184–9.

Wysocka-Diller JW, Helariutta Y, Fukaki H, Malamy JE, Benfey PN. Molecular analysis of SCARECROW function reveals a radial patterning mechanism common to root and shoot. Development. 2000;127:595–603.

# Part V
# Techniques

# Chapter 18
# Forward Genetics in *Setaria viridis*

**Hui Jiang, Pu Huang, and Thomas P. Brutnell**

**Abstract** Forward genetics is a powerful approach to identify mutations and genes underlying traits of interest. Typically, screens begin with chemical mutagenesis of seed or pollen to generate a collection of novel alleles. Dominant or semidominant mutants can be identified in screens of M1 plants and recessive mutants identified in screens of M2 families following self-pollination of the M1. During the last few years, the low cost of next-generation sequencing (NGS) has enabled mapping by sequencing, greatly accelerating the process of gene discovery. As an alternative to mutagenesis, quantitative trait loci (QTL) mapping and genome-wide association studies (GWAS) are powerful tools to explore natural variation. In this chapter, we describe the development of a *Setaria viridis* NMU-mutagenized population using the sequenced reference line A10.1. Strategies for screening mutant populations are described as techniques using bulked segregant analysis by sequencing or direct sequencing of the mutant lines to identify causative lesions underlying mutant phenotypes. We will also discuss strategies to develop recombinant inbred lines (RILs) and MAGIC mapping populations in *Setaria viridis*.

**Keywords** *Setaria viridis* • Forward genetics • NMU mutagenesis • Recombinant inbred line • Bulked segregant analysis by sequencing

## 18.1 Introduction

Screens of mutagenized populations for a phenotype of interest followed by fine mapping to a genetic interval remain a powerful technique for gene discovery in plants. Generally referred to as forward genetics, this approach has been used extensively in many plant model species and numerous crops (Sikora et al. 2011). Although the mutagen may vary (chemical, radiation, or biological such as insertional

H. Jiang • P. Huang • T.P. Brutnell (✉)
Donald Danforth Plant Science Center, 975 North Warson Rd., St. Louis, MO 63132, USA
e-mail: hjiang@danforthcenter.org; phuang@danforthcenter.org; tbrutnell@danforthcenter.org



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_18

mutagenesis by transposable elements or T-DNA insertions), the goal is usually to map a trait of interest to a genetic locus. Validation of the mapping can be achieved either by fine mapping multiple alleles or, when transformation is possible, by complementing the mutant phenotype with a transgene. To explore extant genetic variation in natural populations, the techniques of QTL mapping and Genome-Wide Association Study (GWAS) are often employed. In QTL mapping biparental crosses are performed to generate an F1 hybrid. Self-pollination or backcrossing is then performed to create structured populations that can be screened for variation in traits of interest. In GWAS associations are made between phenotypic and genotypic variation present in large populations often of 200–1000 genetically diverse individuals. Each of these methods offers unique advantages and limitations to exploring genetic variation. In this chapter, we will discuss several methodologies that are under development to facilitate gene discovery in *Setaria viridis* through forward genetic approaches.

## 18.2 Chemical Mutagenesis in *Setaria viridis*

Chemical mutagenesis is an efficient way to create point mutations distributed throughout the genome that produce abundant phenotypic variation and enable the validation of gene function. Compared to other methods of mutagenesis, such as irradiation and insertional mutagenesis, chemical mutagenesis has two major advantages. The first is that it can deliver a uniform density of mutants across the genome, which, depending upon the dosage of mutagen, permits saturation of mutagenesis using relatively few plants. The second advantage is that chemical mutagenesis not only generates loss-of-function mutants but also generates dominant or semidominant gain-of-function alleles. The disadvantage of chemical mutagenesis has traditionally been in the time-consuming effort to clone the gene underlying the mutant phenotype, which requires multiple crosses, fine mapping and positional cloning, and the validation of gene function by genetic transformation. However, with the advent of next-generation sequencing (NGS), the identification of causative mutations has become less rate limiting [e.g., mapping by sequencing (Schneeberger 2014)].

There are three chemical mutagens that have been widely used for mutagenesis: Ethyl methane sulfonate (EMS), *N*-Nitroso-*N*-methylurea (NMU), and *N*-ethyl-*N*-nitrosourea (ENU). As alkylating agents, these mutagens typically cause single nucleotide substitutions and ultimately result in missense, nonsense, or inappropriately spliced transcripts. These three mutagens have been widely used in forward genetics of plants including Arabidopsis, rice, sorghum, maize, barley, wheat, tomato, soybean, medicago, pea , rapeseed, cabbage, melon, oat, peanut, and sunflower (Sikora et al. 2011). EMS can alkylate guanine bases, which results primarily in G/C to A/T transition mutations (Koornneef et al. 1982); NMU functions by transferring its methyl group to nucleic acids, which can lead to AT to GC transition mutations, and has been used for mutagenesis of rice and soybean (Suzuki et al.

2008; Cooper et al. 2008). ENU preferentially induces A to T base transversions, AT to GC transitions, and has also shown to cause GC to AT transitions. ENU has been widely used in mutagenesis of mice (Tokunaga et al. 2014; Moresco et al. 2013) but not extensively in plants.

*Setaria viridis* is an emerging model system for C4 grasses, and closely related to important crops including maize and sugarcane. Therefore, establishing a forward genetics pipeline in *Setaria viridis* can speed the identification and characterization of genes of interest that can be translated to closely related crops. However, to date, no large-scale mutagenesis program has been reported for *Setaria viridis*. We have piloted both EMS and NMU mutagenesis using *Setaria viridis* (A10.1) seeds. The EMS treatment was not effective based on the germination rates of treated seeds and the frequency of sectors in the leaves of M1 plants (data not shown). NMU mutagenesis of *Setaria viridis* was successful and resulted in the development of a population consisting of 20,000 M2 families. The following section describes the methodology employed to achieve the saturation mutagenesis.

## *18.2.1 Materials*

### 18.2.1.1 Choice of Parental Genotype

We chose to conduct mutagenesis in the sequenced accession A10.1 to facilitate downstream analyses of mutants. In Arabidopsis, genotypes targeted for mutagenesis have included 'wild-types' Landsberg erecta (Ler) and Columbia (Col) (Maple and Moller 2007). *Setaria viridis* A10.1 is the current genetic model and has been widely distributed in the *Setaria viridis* research community. *Setaria viridis* A10.1 was also targeted as the reference genome for *Setaria viridis* (recently released 9/20/15) and therefore was the logical initial target for NMU mutagenesis. However, whole genome resequencing has been completed for about 460 *Setaria viridis* accessions at 30× coverage through JGI-DOE (see Chap. 3). Thus, the availability of genome sequence for *Setaria viridis* should not limit the examination of accessions with unique features. For example, it will be beneficial to perform the mutagenesis in the accession ME034V-1, which can be transformed at a higher efficiency (Chap. 20).

### 18.2.1.2 The Dosage of Mutagen

In general, the use of high dosages of chemical mutagens can yield more mutants in a single screen, which allows the screening of relatively small mutagenized populations that contain large numbers of mutations. However, a very high mutational load can cause sterility and loss of viability. Moreover, it can lead to an increase in unwanted mutations at multiple loci, which can later complicate the

downstream positional cloning process. At the other extreme, a mild mutagenic treatment will result in low frequencies of mutations and will therefore require extensive screening to obtain mutants of interest.

A simple method to measure the effectiveness of a mutagenesis experiment is to score the germination rate of the M1 seeds and estimate the frequency of chlorotic sectors in M1 plants. In *Arabidopsis*, a frequency of 0.1–1 % chlorotic sectors in M1 plants can indicate effective mutagenesis, while in the M2 screen, 2–10 % of M2 families should have pigment phenotypes (Maple and Moller 2007). In *Setaria viridis* A10.1, soaking seeds in freshly prepared 20 mM NMU for 3 h resulted in a germination rate of 32 % in M1 seeds, and 17 % of M1 plants had chlorotic sectors; while a treatment of 20 mM NMU for 4 h resulted in a germination rate of 17 % in M1 seeds, and 32 % of the M1 plants had sectors. Therefore, 20 mM NMU for 3 h or 4 h are the optimal conditions to develop the *Setaria viridis* A10.1 mutagenized population. The purity of the NMU can also affect the efficiency of the mutagenesis, and thus the condition of treatment may need to be further optimized depending on distributors of NMU. The NMU tested in earlier experiments was ordered from Sigma (St. Louis, MO).

#### 18.2.1.3 Other Materials for Seed Treatment

- NMU (*N*-nitroso-*N*-methylurea): NMU should be freshly ordered for every experiment because of product instability. Always handle NMU under a chemical fume hood. Wear a lab coat, safety glasses, and doubled gloves for any manipulation. Anything that has been in contact with the NMU should not be taken out from the hood, except in a properly sealed waste container. NMU is an extremely toxic compound when in contact or inhaled. A material safety data sheet (MSDS) must be obtained from the NMU manufacturer and read before conducting the mutagenesis experiment to ensure that all local regulations are met.
- Milli Q water
- Falcon tubes (50 mL)
- Parafilm
- Shaker
- *Setaria viridis* A10.1 seeds with a good germination rate (>90 %)

### *18.2.2 Methods*

#### 18.2.2.1 Generation of the NMU Mutant Population

In a typical experiment, seeds of wild type (M0 seeds) are treated with the mutagen (Fig. 18.1). The plants grown from the mutagenized seeds represent the M1 generation and will be heterogeneous and chimeric for induced mutations. The

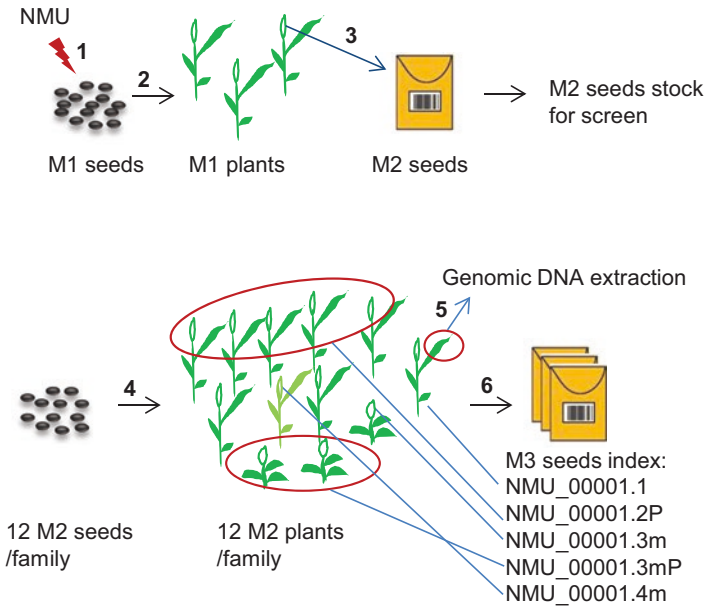


**Fig. 18.1** Schematic of NMU mutagenesis and M2 mutant screen: (*1*) *Setaria viridis* seeds are mutagenized with NMU to produce M1 seeds. (*2*) M1 seeds are grown to produce M1 plants. (*3*) M1 plants are self-pollinated to generate M2 seeds. (*4*) Twelve M2 seeds per family are planted to produce M2 plants. (*5*) Leaf tissue is collected from one individual of each M2 family for genomic DNA extraction. (*6*) M2 generation is self-pollinated to produce M3 seeds. M3 Seeds harvested from different individuals in the same M2 family are indexed according to their phenotype and are stored at the seed bank for further verification. *P* pool, *m* mutant

progeny derived from the self-pollinated M1 plants are the M2 generation, in which homozygous recessive mutations can be detected, and mutant phenotypes will be segregating in each family. Therefore, the M2 generation is most commonly used for forward genetic screens. However, it is important to note, that because the M1 plants are chimeras, M2 families will not always display a 3:1 segregation pattern for single recessive alleles. Thus, M3 populations are often used to examine patterns of segregation and define the dominance or recessive nature of induced alleles.

Seeds of *Setaria viridis* A10.1 were mutagenized with different concentrations of NMU and variable times of treatment. One day prior to the experiment, 4000 seeds were placed on Petri dishes (12 cm) containing Whatman filter paper imbibed with MilliQ (MQ) water, which facilitates the uptake of NMU (approximately 1000 seeds/petri dish). The seeds were then stored at 4 °C for 24 h. On the day of NMU treatment, seeds were transferred from petri dishes into 50 mL Falcon tubes, approximately 2000 seeds/tube. For the stock solution (100 mg/mL), NMU (powder) was dissolved by adding MQ water and vortexing for at least 2 min to ensure that all NMU powder was dissolved completely. The stock solution was then diluted to the working concentration at 20 mM (i.e.,

60 mg/30 mL, or 2 mg/mL; mol wt. = 103.8 mg/mol) by adding MQ water to make the desired volume for all seeds. About 30 ml NMU working solution was added into each Falcon tube that contained 2000 seeds. The Falcon tubes were sealed tightly with Parafilm and then placed horizontally in a shaker (200 rpm at room temperature) for 3 h and 4 h, respectively. After shaking, the NMU solution was poured out in an appropriate container (for decontamination). The seeds were then rinsed, at least four times, with MQ water. All the waste should be stored in a safe area in the hood prior to deactivation. Once rinsed, the seeds were placed on a paper towel and left to dry for 30 min before sowing. Seeds can be sown directly in soil at a density of approximately 200–300 seeds/flat (each flat containing 18 cells). At the same time, nontreated A10.1 seeds should be sown in one tray with the same density to be used as a control for germination and phenotype scoring. The tray should be bottom watered to keep the soil moisture as described (Jiang et al. 2013). M1 plants were grown in a growth chamber (31 °C/22 °C (day/night), 12 h light/12 h dark, RH: 30–50 %), or in a green house (31 °C/22 °C (day/night), 16 h light/8 h dark, RH: 30–50 %). The germination rate and the frequency of white/yellow sectors in leaves of M1 plants were scored and used to estimate the effectiveness of NMU treatment. In general, A10.1 plants flower 21–23 days after sowing when grown in the growth chamber (under short day conditions), and 28–30 days after sowing when grown in the green house (under long day conditions). A10.1 seeds tend to shatter once becoming mature; therefore, it is important to decide the right time for seed harvesting to avoid losing seeds from M1 plants due to shattering. In general, seeds should be harvested 3–4 weeks after flowering.

#### 18.2.2.2 Strategy for M2 Seed Harvesting

We have employed two strategies for harvesting seed of *Setaria viridis* following mutagenesis. The first is to harvest M2 seeds from each M1 plant. Although it involves substantially more effort and time, this method is advantageous when two mutants with a similar phenotype are recovered from two different M1 families, as they are considered two independent mutations. In addition, the dominant/recessive nature of a mutant phenotype can be approximated based on the segregating ratio in each M2 family, facilitating downstream analyses. The second strategy is to pool the M2 seeds from M1 plants grown in each tray, approximately 200 M1 individuals/pool. The advantage of pooling is that it saves time and effort for seed harvesting. However, the disadvantage of this method is that if M2 plants from the same tray of M1 individuals have the same mutant phenotype one should assume that the mutants have been derived from the same M1 plant. Therefore, only plants from different pools are considered independent mutants. Both strategies have been used in the development of our mutant collections. Approximately 15,000 M2 families were harvested from individual M1 plants and 5000 M1 plants were harvested as pools (~200 M1 plants/pool).

M2 seeds from 1 to 2 panicles per M1 individual were harvested and stored in barcoded indexed coin envelopes (Fig. 18.1). Seeds were dried at 30 °C for 3 days in a seed drier. The dried seeds were then stored in the laboratory (T: 24.1 °C ± 0.13 °C, RH: 21.2 ± 1.15 %) or in a seed chamber (T: 4.0 °C ± 1.0 °C, RH: 20 % ± 1 %) for future screening.

#### 18.2.2.3 Protocol for Screening M2 Population

In order to maximize the number of recovered mutants presenting phenotypes of interest in each M1 plant, 12 seeds of each NMU M2 family are planted and scored for mutant phenotypes. As mentioned earlier, screens of M3 families are conducted to determine the dominant/recessive nature of the mutant alleles. When screening M2 families, a few *Setaria viridis* A10.1 seeds should also be planted in each flat as a control for phenotypic evaluation. To survey phenotypes in the population, approximately 2000 M2 families were grown in the green house (31 °C/22 °C day/night, 16 h light/8 h dark, RH: 30–50 %) and plants scored at multiple developmental stages to identify mutant phenotypes. At 7–10 days after planting, albino, pale green, and yellow leaf mutants can be scored but mutants die soon after seed reserves are depleted. At 2–5 weeks after planting, leaf phenotypes including zebra stripe, tie dye, rough narrow leaf, and adherent leaf can be identified along with plant stature mutants, including dwarf, stunted, and small. Other phenotypic traits that were examined included days to panicle emergence (PE), tiller number, and panicle mutants. Once an interesting mutant phenotype is identified, the individual that displays the mutant phenotype is flagged and bagged for seed harvesting and further verification. In general, 5–6 weeks after planting, all plants should be bagged individually or as a pool according to their identity to avoid seed lost. Usually seeds can be harvested 6–8 weeks after planting. Also, mutant phenotypes including low/nonseed shattering and late senescence can be scored at seed harvesting.

#### 18.2.2.4 Diversity of Mutant Phenotypes

Of the approximately 2000 M2 families screened, approximately 30 % displayed visible mutant phenotypes. A diversity of phenotypes related to plant development and morphology were observed including abnormal leaf morphology and pigmentation (Fig. 18.2), plant stature and flowering time variants (Fig. 18.3), panicle phenotypes (Fig. 18.4), and seed size mutants (Fig. 18.5). All the aforementioned mutants present morphological changes that are easily identified. However, the development of high-throughput phenotyping methods for screening will enable more sophisticated screens that could aid in the identification of phenotypes that are not easily observed by eye such as alterations in growth rate, altered photosynthetic efficiency (e.g., Fv/Fm), and altered water use efficiencies.

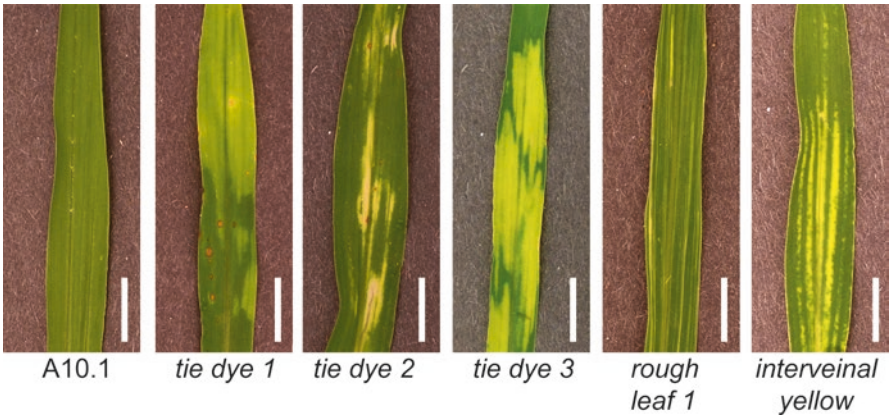


**Fig. 18.2** Representative mutant leaf phenotypes of *Setaria viridis* NMU mutants. Scale bar: 1 cm

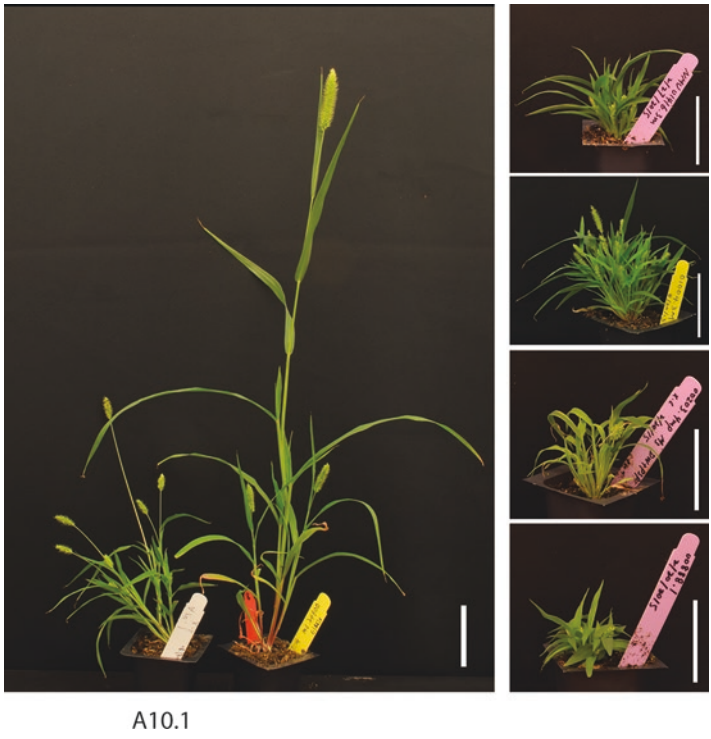


**Fig. 18.3** *Setaria viridis* mutants with altered plant stature phenotypes. *Left panel* shows the comparison between wild-type A10.1 and a mutant with large panicle, tall stature, and late flowering phenotype. *Right panel* shows four mutants with short stature phenotype. Scale bar: 5 cm

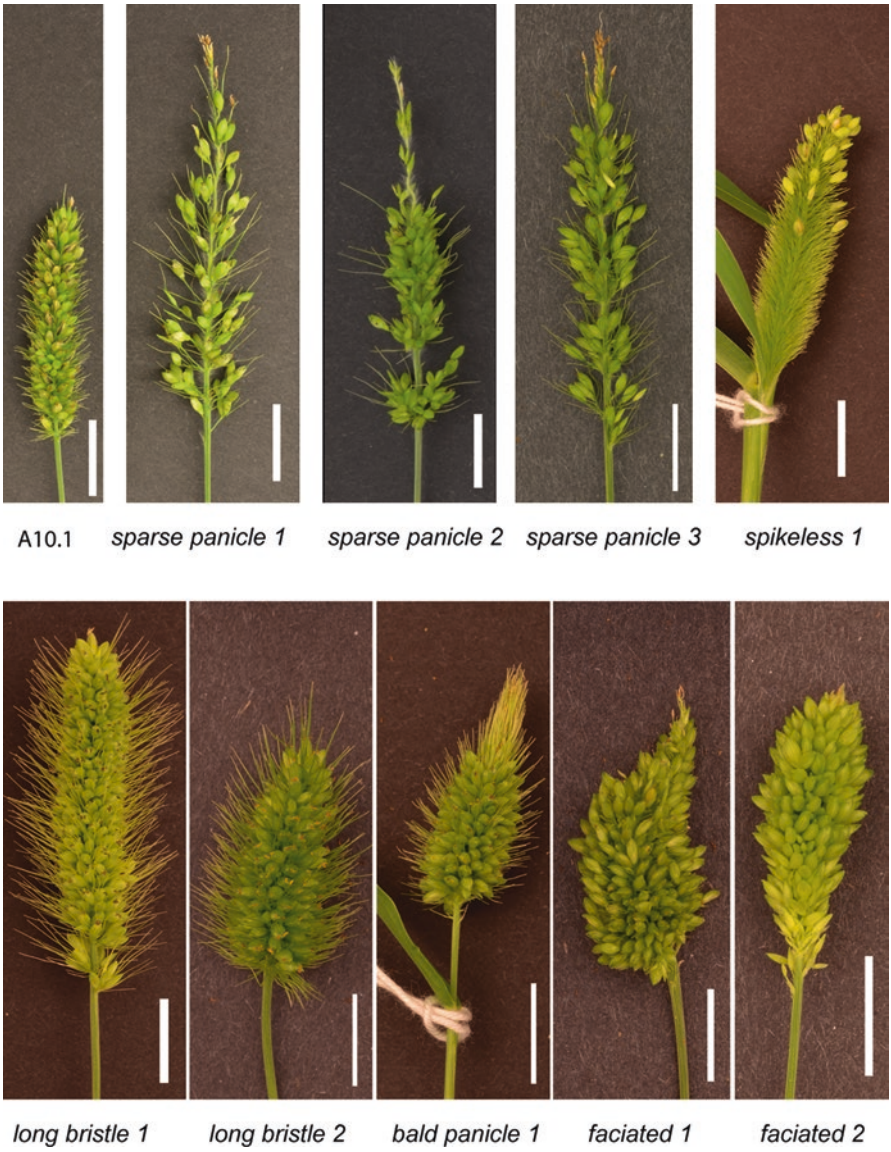


**Fig. 18.4** Panicle phenotypes of NMU mutants. Bar length: 1 cm

#### 18.2.2.5 Indexing of M3 Seeds

To maintain the integrity of the M3 populations and to track the seed pedigree, M3 seeds harvested from individual plants of the M2 family NMU_00001 were indexed (Table 18.1).

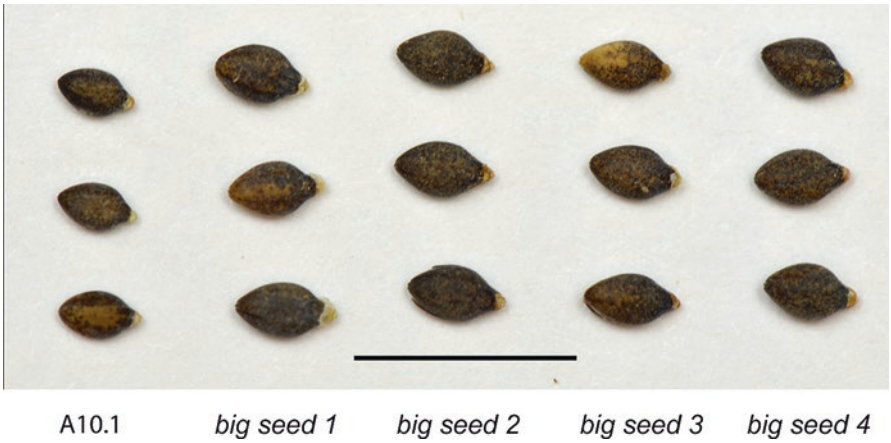


**Fig. 18.5** Seed size phenotypes of NMU mutants. Three seeds from each NMU family are used for comparison. Scale bar: 5 mm

**Table 18.1** Index of M3 seeds from a M2 family NMU_00001

| Index of M3 seeds | Description |
|---|---|
| NMU_00001.1 | Seeds of one individual plant from which genomic DNA was extracted |
| NMU_00001.2P | Seeds of a pool of individual plants that do not show an obvious phenotype |
| NMU_00001.3m | Seeds of one individual plant that presents a phenotype of interest |
| NMU_00001.3mP | Seeds of a pool of individual plants that show the same mutant phenotype as that of individual NMU_00001.3m |
| NMU_00001.4m | Seeds of one individual plant that has a different phenotype of interest from that of NMU_00001.3m |

*P* pool, *m* mutant

### *18.2.3 Bulked Segregant Analysis (BSA) by Sequencing in* Setaria viridis

The advent of next-generation sequencing technologies has enabled the development of new strategies to identify the causative lesion underlying mutants of interest (Schneeberger 2014). As detailed later, we have developed an approach to fine map genes of interest using a modified bulked segregant approach (Michelmore et al. 1991).

#### 18.2.3.1 Experimental Design: Crossing Scheme, Pool Size, and Coverage

The experimental design of fine mapping depends on many aspects, including the genome size of the organism, generation time, feasibility of crossing, penetrance and dominance of the mutation, mutation rates, as well as total cost to map the causative

mutation. If possible, mutants should be self-pollinated to generate a true-breeding stock for phenotypic evaluations and outcrossed to the parental line or genotypically diverse line for bulked segregant analysis.

One advantage to crossing the mutant with a nonparental line is that the existing natural variation and induced chemical variation may serve as genetic markers, resulting in more precise candidate gene intervals, although it is dependent on the number of recombination events in the F2 mutant pool. The large number of markers distributed throughout the genome enables the accurate assessment of allele frequencies even at low sequence coverage levels. This is particularly useful when the mutation frequency of the mutagenized lines is low resulting in few markers for fine mapping. In *Arabidopsis*, sequencing a pool of 500 mutant F2 recombinants at a 22× genome coverage resulted in the identification of a single nonsynonymous SNP candidate, and later confirmed as the causative SNP for slow growth and light green leaves mutant phenotype (Schneeberger et al. 2009). In *Mimulus lewisii*, Yuan et al. (2013) pooled 100 F2 mutant recombinants of an outcrossed population and sequenced it at 55× coverage of the pool. In this study, it was possible to fine map the *guideless* mutant within a 50 kb interval that contained nine genes, in which a 2-bp frameshift insertion in an exon was confirmed as the causative mutation for the change of formation of nectar guides. However, one major disadvantage of outcrossing populations is that it introduces phenotypic variation that can interfere with subtle phenotypes, depending on the penetrance and expressivity of the mutant phenotype. Scoring complex and subtle phenotypes increases the likelihood of inconsistent or inaccurate categorization of plants, which may have severe effects on mapping by sequencing, especially when the pool size of F2 mutants is relatively small (Schneeberger 2014).

In contrast, backcrossing the mutant to the parental line allows for subtle and complex phenotypes to be scored without confounding second site suppressors or enhancers. Another advantage is that only SNPs generated by mutagenesis are used as markers to map the candidate region harboring the causative mutation. Thus, fewer SNPs need to be considered, and the causative mutation can be easily identified if the coverage is sufficient. Moreover, fewer F2 mutant individuals are needed to bulk sequence and map the causative mutation compared to the outcrossing population (Abe et al. 2012).

Ultimately, the final number of candidate causal mutations is strongly affected by the sequencing coverage and the mutation rate of the mutagenesis (Schneeberger 2014; James et al. 2013). An in silico study in Arabidopsis showed that mapping in near-isogenic populations required higher sequence coverage compared to mapping in diverse populations (James et al. 2013). An empirical study in Arabidopsis showed that one causal mutation was identified for a leaf hyponasty mutant after sequencing a pool of 110 backcrossed F2 (BCF2) mutants at 50× coverage (Allen et al. 2013). Austin et al. (2011) reported that one gene involved in cell wall biology in Arabidopsis was successfully mapped using as few as ten BCF2 progeny. In rice, Abe et al. (2012) and Takagi et al. (2015) sequenced two pools of 20 BCF2 mutant progeny at a coverage >12×, and successfully mapped causative mutations responsible for pale green leaf and *hitomebore salt tolerant 1(hst1)* mutant phenotypes, respectively.

*Setaria viridis* is similar to *Oryza sativa*, in terms of genome size, outcrossing rate (less than 1 %), and residual heterozygosity in the genome. Thus, the MutMap method for bulked-segregant analysis by whole genome resequencing of a BCF2 progeny used in rice (Abe et al. 2012) may serve as a useful model for mapping genes in *Setaria viridis*. A major advantage in developing an isogenic backcross population in *Setaria viridis* is that an optimal method for crossing A10.1 has been established (Jiang et al. 2013), and the current NMU mutant population has been developed in the A10.1 background. In this case, no further optimization for crossing is needed if A10.1 is used as the female parent. In summary, to map an NMU mutant in *Setaria viridis*, it is recommended that mutant plants be backcrossed to the parental line and a pool of >20 F2 mutant recombinants be used for deep sequencing. Sequence data for these lines at 20×–30× coverage (at least 1× coverage/mutant) should narrow the candidate gene interval to a small number of SNP candidates. To demonstrate the protocol for bullked-segregant analysis by sequencing in *S. viridis*, an example is shown later for the *sparse panicle 2* mutant (*ssp2*).

#### 18.2.3.2 Protocol of Bulked-Segregant Analysis by Sequencing in *S. viridis*

A recessive mutant with a sparse panicle phenotype (named as *spp2*) was identified in a NMU M2 screen and was then self-pollinated to produce M3 seed fixed for the mutant phenotype. To map the gene that underlies the sparse panicle phenotype, the mutant was crossed as a male onto A10.1, the wild-type progenitor. Three F1 progeny were planted and all F1 plants displayed a wild-type panicle phenotype, indicating that the *spp2-1* mutant allele is recessive (Fig. 18.6). This was confirmed after scoring 82 F2 progeny that segregated wild type:sparse panicle in an approximately 3:1 ratio (58:24,chi sq. *p* value=0.3721), indicating that the *ssp2-1* allele is recessive. Leaf samples from each individual presenting the sparse panicle phenotype were collected for DNA extraction using a modified CTAB method (http://www.danforthcenter.org/scientists-research/principal-investigators/thomas-brutnell/resources). Each DNA sample was treated with RNaseA, and the final DNA concentration determined using the Qubit broad DNA quantification kit (Lot No. 1691775, Invitrogen). Equal quantities of DNA from a total of 20 mutants were pooled. Pooled DNA was sheared and used for Illumina library construction following the methods of Lindner et al. (2012) and sequenced to a depth of 30× coverage. The MutMap method (Abe et al. 2012) for sequence analysis was used to identify the causative mutation for the *ssp2* mutant phenotype. The general principle of the MutMap method is to first scan the genome to create a SNP index, defined as the ratio between the number of reads of a mutant SNP and the total number of reads corresponding to the SNP interval, in order to find the region with a SNP index of 1. This region is expected to harbor the gene responsible for the mutant phenotype, because in a pool of BCF2 mutants, the causative SNP and the closely linked SNP should be 100 % mutant and 0 % wild type for all of the reads. Using this method, we anticipate delimiting the mutant interval to approximately 1 Mb, which may

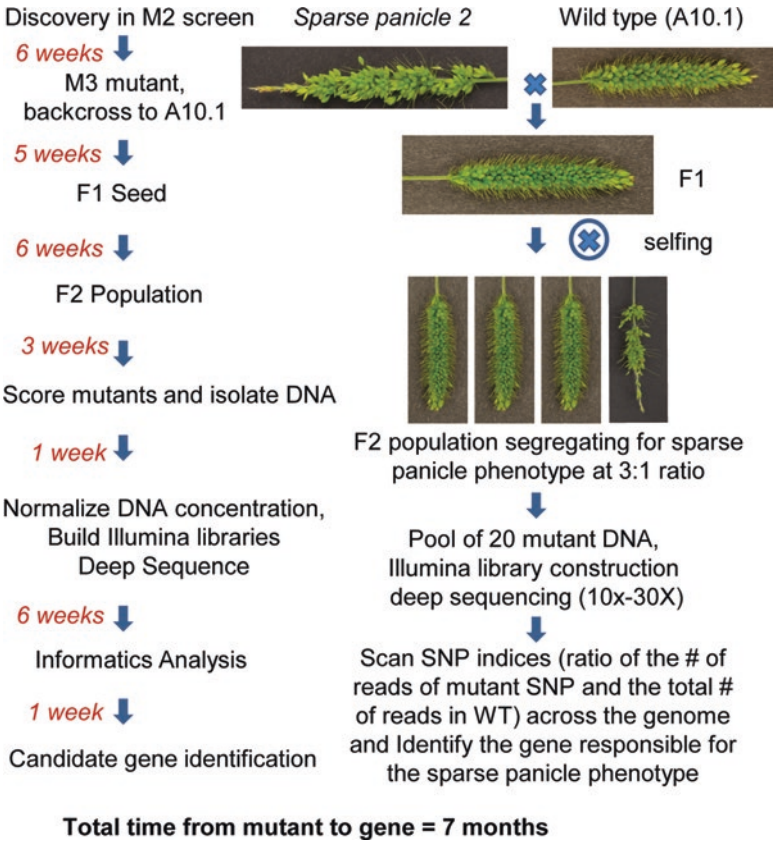


**Fig. 18.6** Schematic of bulked segregant analysis by sequencing in *S. viridis. Right panel*: procedure for bulked segregant anlaysis by sequencing to map a *sparse panicle 2* mutant. *Left panel*: time line from mutant discovery to candidate gene identification

contain about five disruptive SNPs, thus a short list of candidate genes can be generated based on gene annotation. Once the genes are identified, they can be validated using CRISPR/Cas9 technology (see Chap. 19).

### 18.2.3.3 Applications and Perspective

The subfamily Panicoideae contains some of the world's most productive cereal grasses, including maize and sorghum. However, the majority of these grasses are large and long lived, which impedes efforts to identify and validate gene candidates. *Setaria viridis* is a tractable model for rapid gene candidate identification due to its short life cycle and small stature. As shown in (Fig. 18.6), we estimate that the time from mutant discovery to candidate gene can be as short as 7 months. In contrast, the same approach in maize would take at least 14 months in a greenhouse or longer

in the field. Furthermore, only six flats (18 inserts/flat) are required to screen 100 *S. viridis* M2 families (12 individuals/family) while an entire greenhouse is needed to screen 100 maize individuals to maturity. Moreover, in order to achieve the same genome sequence coverage, the cost of sequencing a bulked BCF2 population in maize will be four times that of *Setaria*, since the maize genome size (~2300 Mb) is four times larger than *S. viridis* (~515 Mb). In conclusion, the use of *S. viridis* for BSA by sequencing to map genes that underlying traits of interest has the advantages of time, labor, and cost saving.

## *18.2.4 Direct Sequencing of Mutants and Mutant Families*

### 18.2.4.1 Why Direct Sequencing

Sequencing of the F2 mutant pool by BSA is a reliable approach to map the causal mutation(s) of a mutant phenotype. On the other hand, direct sequencing of individuals or family pools with mutant phenotypes is becoming an increasingly feasible approach. The biggest advantage of direct sequencing compared to BSA is that it is both simpler in terms of labor and faster. As mentioned earlier, BSA requires two generations (a backcross and self-pollination), DNA extraction, and pooling of a large number of F2 mutants. When the mutant phenotype is not directly visible (e.g., photosynthetic rate), screening a large number of F2 lines for a mutant phenotype can be quite labor intensive. With the BSA approach it takes approximately 7 months to identify the causal gene/region in *S. viridis*, and much longer (some times more than 14 months) in crop species such as maize and sorghum. Direct sequencing of an individual or a pool of mutant families can be done within weeks, depending on the turnaround time of the sequencing results.

Mutagenesis by NMU usually generates thousands to tens of thousands of SNPs in the genome (Koornneef et al. 1982). A major challenge in the direct sequencing approach is that it might not provide enough resolution to map the causal mutation down to a single SNP or gene interval, in comparison to BSA. However, the results from direct sequencing may be reasonably satisfactory if coupled to other approaches to prioritize the list of genes and candidates for targeted approaches. Targeted approaches such as RNAi or with gene editing are becoming increasingly easier, faster, and precise, and are also generally required to fully characterize the functions of a target gene. Following we discuss additional methods to prioritize candidate genes.

### 18.2.4.2 Prioritizing Candidate Genes from Direct Sequencing

***Homozygosity***. First, to observe a fixed stable mutant phenotype, the locus of the causal mutation of the mutant individual or mutant family should be homozygous, assuming it is a recessive mutant phenotype. A plot can then be generated for the alternative allele count for every SNP observed in the genome versus the physical

location where the SNP is mapped to determine regions with homozygosity. The causal mutation is most likely contained in these regions. If the mutant individual/family pool is in the M3 generation, this can exclude about half of the genome. On average, successive backcrossing to a wild-type progenitor will further reduce the number of potential causative SNPs but will not save time over BSA.

***SNP effect prediction***. The causal SNP is generally expected to have a large phenotypic effect. Software packages such as snpEff (Cingolani et al. 2012) can categorize SNPs into different functional categories—intergenic, intronic, synonymous, missense, gain/loss of intron splicing sites, and gain/loss of start/stop codon. The first three categories, intronic, intergenic, and synonymous SNPs are less prone to generate phenotypic effects; while the others (denoted as disruptive), especially premature stop codons and the loss of a start codon, are more likely to result in a loss-of-function allele, and thus have a higher chance of inducing an obvious mutant phenotype. For missense mutations, the effects of amino acid substitutions can be potentially predicted via homology-based searches using packages such as SIFT (Ng and Henikoff 2001). For example, a change of amino acid in a highly conserved motif relating to a particular molecular function is more likely to cause strong allelic effects.

***Gene annotations***. Another intuitive way to identify causal mutations is to examine the gene annotations directly. Similar molecular functions and phenotypic effects may be expected between *S. viridis* and other grasses, or even in dicot species such as *Arabidopsis*. The existing annotations of the list of candidate genes can be further refined using tools like Blast2Go (Conesa et al. 2005). Although this criterion involves more manual inspection, examining the gene annotations occasionally can identify some "obvious candidates" depending on the mutant phenotypes: for example, a decreased tillering phenotype with a mutation in the TEOSINTE BRANCHED 1 ortholog.

***Summarizing criteria and limitations***. None of the criteria mentioned earlier can alone satisfactorily make predictions of candidate causal mutations. For example, SNP effect predictions may ignore SNPs in the promoter region leading to expression level changes, while gene annotations alone may bring a large number of seemingly related candidates. Both methods are also affected by misannotation in grass genomes in general. However, summarizing information from both methods could lead to a higher chance of predicting the causal SNP (e.g., a homozygous SNP with a premature stop codon that excludes an important functional domain in a gene that potentially is correlated to the phenotype). Additional information, such as coexpression gene networks and a gene atlas for *Setaria* will be available soon (Jiang & JGI unpublished data) and should also be integrated into the analyses. Ultimately, when sufficient experimental data is accumulated, comprehensive predictive models can be built to make more reliable predictions. Importantly, these models can be applied to BSA results as well (when BSA mapping resolution is not enough to identify a single gene).

When multiple mutants of similar phenotype are available, there is the possibility to fine map using direct sequencing. For instance, if two mutants are allelic (i.e., different mutations in the same underlying gene), then the causal mutation should be within homozygous regions in both mutant families, and likely have independent nonsynonymous changes in the two mutant families. This can greatly narrow down the list of candidate genes. For instance, in two mutant families (*grainy leaf 1*, and *grainy leaf 2*) with similar phenotype (virescent or pale), we identified only one single gene with independent homozygous nonsynonymous mutations in both mutant families, making it a likely candidate. We are currently in the process of further dissecting the mutant phenotype and testing allelic status by crossing the two mutant families.

#### 18.2.4.3 Other Usages of Direct Sequencing

In addition to directly prioritizing a list of candidate genes, direct sequencing of individual mutants or pools of mutants from the same family can be useful in other ways. First, it gives estimates of mutation frequency and disruptive mutation frequency in the mutant population. This provides a base line to calculate the size of a mutant population that is required to saturate the genome with certain number of high impact SNPs. Second, it can provide markers to screen F1 hybrids. When a wild type is used as the maternal parent (e.g., low fertility or difficulties in performing crosses using mutant lines as a female parent), a hybrid is not directly visible and molecular makers are required to identify a true F1 hybrid. Third, it provides markers for a more precise BSA. By examining the mutant allele frequency in the sequenced mutant lines, SNPs caused by mutagenesis should have an allele count of 1 (heterozygotes) or 2 (homozygotes). Thus, they can be distinguished from remaining heterozygotes or SNPs caused by mapping errors that most likely have a frequency higher than 2. Finally, it could potentially facilitate reverse genetics. If a gene of interest is known to have a strong effect in a mutant individual, predicted phenotypes can be further examined in that individual, especially when the phenotypes are difficult to screen using forward genetic approaches.

## 18.3 Development of Recombinant Inbred Lines (RILs) in *S. viridis*

### *18.3.1 Introduction*

During the last few decades, there have been many studies in plant model species and crops that exploit natural variation as a forward genetics tool. The use of recombinant inbred lines (RILs) is important to dissect the genetic architecture of complex traits, since each line is nearly homozygous and can be propagated as genetically identical individuals. RIL populations thus allow the genotyping and phenotyping of many traits under various environmental conditions. In Arabidopsis,

around 60 RIL populations are available from stock centers (Weigel 2012), and methods have been developed to speed up the production of double haploid plants from RILs (Ravi and Chan 2010). All these resources have greatly facilitated the study of natural variation in Arabidopsis and allowed fundamental questions of biology to be addressed. The classical biparental RIL population provides high mapping power but is limited by low resolution due to the shortage of diversity and recombination events. Recently, multiparent advanced generation intercross (MAGIC) populations have been developed in many model plants and crops, including Arabidopsis, rice, wheat, and maize (Kover et al. 2009; Bandillo et al. 2013; Mackay et al. 2014; Dell'Acqua et al. 2015; Huang et al. 2015). MAGIC populations have produced populations with abundant genetic diversity but less subpopulation structure than association mapping, and higher mapping power than biparental mapping panels (Dell'Acqua et al. 2015). MAGIC populations are thus an important resource for dissecting the genetic basis of natural variation. In rice, MAGIC populations have provided important prebreeding materials harboring useful traits derived from multiple elite breeding lines.

*Setaria viridis* is an emerging model for C4 grasses, thus identification of genes in *S. viridis* may lead to the identification of homologous loci important for agronomical important traits in important crop and biofuel grasses. However, to date, in Setaria only two RIL populations derived from crosses between *S. viridis* and *S. italica* have been published. One population was derived from a cross between *S. italica* accession B100 and *S. viridis* accession A10 (Wang et al. 1998; Bennetzen et al. 2012), and a number of QTLs for flowering time difference, shattering, height, branching, and biomass in *Setaria* were identified using 182 F7 RILs from this population (Mauro-Herrera et al. 2013; Doust et al. 2014; Mauro-Herrera and Doust 2016) (see also Chap. 12). Another *Setaria* RIL population was derived from a cross between Yugu1, a *S. italica* cultivar developed in China, and a wild *S. viridis* accession collected from Uzbekistan. Several osmotic stress-related QTLs were identified by using 190 F7 lines from the Yugu1 × *S. viridis* population (Qie et al. 2014). To explore the rich genetic diversity among wild *S. viridis* accessions, and to reduce the segregation distortion due to the crosses performed between different subspecies, RIL populations derived from diverse *S. viridis* × *S. viridis* accessions are needed.

Recently, a collaboration between several laboratories and JGI-DOE has enabled whole genome resequence of 430 Setaria accessions at 30× coverage. The genetic diversity and population structure of these largely North American *S. viridis* accessions have been studied (Huang et al. 2014) (Chap. 3). In our laboratory, six *S. viridis* RIL populations have been initiated, derived from crosses made between A10.1, a *S. viridis* accession used as a reference genome sequence (http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sviridis_er), and six diverse *S. viridis* lines, which were selected based on 12 SSR markers to maximize the genetic diversity (from a set of 150 *Setaria viridis* accessions). Among the six RILs, one RIL population consists of 280 RILs derived from a cross between A10.1 and Roche 10106, a drought tolerant accession. This A10.1 × Roche 10106 population has been developed through single seed decent (SSD) to the F5 generation and will be self-pollinated to the F8 generation. Genotyping of F5 progeny will be analyzed

using GBS and F6 seeds will be retained as HIFs (heterogeneous inbred families), which will be used for fine mapping QTLs for trait of interest such as the identification of drought stress related QTLs.

### *18.3.2 Protocol to Develop* S. viridis *RIL Populations*

***Parental line selection****:* There are several possible strategies to identify parental lines that are largely dependent on the goals of the QTL mapping experiment. In all cases, diversity estimates of the populations should first be quantified and estimates of population structure determined. Populations can then be generated from genetically diverse individuals. However, the choice of parental lines will largely vary with goals for mapping. For instance, if the objective is to map genes that contribute to heat and cold tolerance, a population can be generated by crossing a parent line that was isolated from a hot environment (e.g., Texas) to a line isolated from a cool environment (e.g., Minnesota) with the goal of capturing a range of diverse alleles that contribute to growth under temperature extremes. Another approach is to screen a core set of *Setaria viridis* accessions, which represent the maximum genetic, geographic, and environmental diversity for abiotic stress tolerance. Selected tolerant lines can be used as parents to develop RILs in order to identify genes involved in different pathways that plants have developed to adapt to similar environments.

***RIL development****:* Once two parents are selected, they are crossed to each other to develop biparental RIL populations. The F1 seeds are self-pollinated to F2 progeny. Ideally at least 200 F2 progeny should be self-pollinated and true-breeding lines generated through single seed descent (SSD) to the F8 generation. Genotyping of RILs through GBS will enable the determination of breakpoints and aid in the development of markers. To fine map traits of interest we recommend genotyping at the F5 generation using GBS, and to retain the F6 seeds as HIFs, which will be used to fine map the QILs of interest that will be discovered afterward.

### *18.3.3 Future Perspectives*

In addition to developing RIL populations, the use of MAGIC populations (Huang et al. 2015) could complement GWAS studies of natural collections to genetically dissect traits such as tolerance to abiotic stresses such as drought, cold, low nitrogen, or phosphate. Ultimately, a major goal in our group is to understand the mechanisms underlying variation to abiotic stress. Thus, exploiting a number of approaches from mutagenesis to GWAS surveys will enable a deep exploration of the genetic variation that underlies traits of interest. What will facilitate all of these studies is access to well-characterized germplasm. Thus, the storage and distribution of seed

stocks will be critical for the community. Currently, seeds can be accessed in the US through the USDA (http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?430573). Linking genotypic data with specific germplasm will be increasingly important as genome sequencing technologies reduce the time and expense associated with whole genome sequencing.

**Acknowledgement** This work was supported by grants from Department of Energy (DE-SC0008769). The authors thank Hugues Barbier, Xiaoping Li, and Zhonghui Wang for assistance in generating the NMU mutant population, and Carla Coelho and Malia Gehan for helpful comments to the manuscript.

## References

Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, et al. Genome sequencing reveals agronomically important loci in rice using MutMap. Nat Biotechnol. 2012;30(2):174–8. Epub 2012/01/24.

Allen RS, Nakasugi K, Doran RL, Millar AA, Waterhouse PM. Facile mutant identification via a single parental backcross method and application of whole genome sequencing based mapping pipelines. Front Plant Sci. 2013;4:362. Epub 2013/09/26.

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, et al. Next-generation mapping of Arabidopsis genes. Plant J. 2011;67(4):715–25. Epub 2011/04/27.

Bandillo N, Raghavan C, Muyco PA, Sevilla MA, Lobina IT, Dilla-Ermita CJ, et al. Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice (N Y). 2013;6(1):11. Epub 2013/11/28.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, et al. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30(6):555–61.

Cingolani P, Platts A, le Wang L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6(2):80–92. Epub 2012/06/26.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005;21(18):3674–6. Epub 2005/08/06.

Cooper JL, Till BJ, Laport RG, Darlow MC, Kleffner JM, Jamai A, et al. TILLING to detect induced mutations in soybean. BMC Plant Biol. 2008;8:9. Epub 2008/01/26.

Dell'Acqua M, Gatti DM, Pea G, Cattonaro F, Coppens F, Magris G, et al. Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in Zea mays. Genome Biol. 2015;16:167. Epub 2015/09/12.

Doust AN, Lukens L, Olsen KM, Mauro-Herrera M, Meyer A, Rogers K. Beyond the single gene: how epistasis and gene-by-environment effects influence crop domestication. Proc Natl Acad Sci U S A. 2014;111(17):6178–83. Epub 2014/04/23.

Huang P, Feldman M, Schroder S, Bahri BA, Diao X, Zhi H, et al. Population genetics of Setaria viridis, a new model system. Mol Ecol. 2014;23(20):4912–25. Epub 2014/09/05.

Huang BE, Verbyla KL, Verbyla AP, Raghavan C, Singh VK, Gaur P, et al. MAGIC populations in crops: current status and future prospects. Theor Appl Genet. 2015;128(6):999–1017. Epub 2015/04/10.

James GV, Patel V, Nordstrom KJ, Klasen JR, Salome PA, Weigel D, et al. User guide for mapping-by-sequencing in Arabidopsis. Genome Biol. 2013;14(6):R61. Epub 2013/06/19.

Jiang H, Barbier H, Brutnell T. Methods for performing crosses in Setaria viridis, a new model system for the grasses. J Vis Exp. (80); 2013. Epub 2013/10/15.

Koornneef M, Dellaert LW, van der Veen JH. EMS- and radiation-induced mutation frequencies at individual loci in Arabidopsis thaliana (L.) Heynh. Mutat Res. 1982;93(1):109–23.
Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, et al. A multiparent advanced generation inter-cross to fine-map quantitative traits in Arabidopsis thaliana. PLoS Genet. 2009;5(7):e1000551. Epub 2009/07/14.
Lindner H, Raissig MT, Sailer C, Shimosato-Asano H, Bruggmann R, Grossniklaus U. SNP-Ratio Mapping (SRM): identifying lethal alleles and mutations in complex genetic backgrounds by next-generation sequencing. Genetics. 2012;191(4):1381–6. Epub 2012/06/01.
Mackay IJ, Bansept-Basler P, Barber T, Bentley AR, Cockram J, Gosman N, et al. An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation. G3 (Bethesda). 2014;4(9):1603–10. Epub 2014/09/23.
Maple J, Moller SG. Mutagenesis in Arabidopsis. Methods Mol Biol. 2007;362:197–206. Epub 2007/04/10.
Mauro-Herrera M, Doust AN. Development and genetic control of plant architecture and biomass in the panicoid grass, Setaria. PLoS One. 2016;11(3):e0151346. doi:10.1371/journal.pone.0151346.
Mauro-Herrera M, Wang X, Barbier H, Brutnell TP, Devos KM, Doust AN. Genetic control and comparative genomic analysis of flowering time in Setaria (Poaceae). G3 (Bethesda). 2013;3(2):283–95. Epub 2013/02/08.
Michelmore RW, Paran I, Kesseli RV. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci U S A. 1991;88(21):9828–32. Epub 1991/11/01.
Moresco EM, Li X, Beutler B. Going forward with genetics: recent technological advances and forward genetics in mice. Am J Pathol. 2013;182(5):1462–73.
Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001;11(5):863–74. Epub 2001/05/05.
Qie L, Jia G, Zhang W, Schnable J, Shang Z, Li W, et al. Mapping of quantitative trait locus (QTLs) that contribute to germination and early seedling drought tolerance in the interspecific cross Setaria italica × Setaria viridis. PLoS One. 2014;9(7):e101868. Epub 2014/07/18.
Ravi M, Chan SW. Haploid plants produced by centromere-mediated genome elimination. Nature. 2010;464(7288):615–8. Epub 2010/03/26.
Schneeberger K. Using next-generation sequencing to isolate mutant genes from forward genetic screens. Nat Rev Genet. 2014;15(10):662–76. Epub 2014/08/21.
Schneeberger K, Ossowski S, Lanz C, Juul T, Petersen AH, Nielsen KL, et al. SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat Methods. 2009;6(8):550–1. Epub 2009/08/01.
Sikora P, Chawade A, Larsson M, Olsson J, Olsson O. Mutagenesis as a tool in plant genetics, functional genomics, and breeding. Int J Plant genomics. 2011;2011:314829. Epub 2012/02/09.
Suzuki T, Eiguchi M, Kumamaru T, Satoh H, Matsusaka H, Moriguchi K, et al. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. Mol Genet Genomics. 2008;279(3):213–23.
Takagi H, Tamiru M, Abe A, Yoshida K, Uemura A, Yaegashi H, et al. MutMap accelerates breeding of a salt-tolerant rice cultivar. Nat Biotechnol. 2015;33(5):445–9. Epub 2015/03/24.
Tokunaga M, Kokubu C, Maeda Y, Sese J, Horie K, Sugimoto N, et al. Simulation and estimation of gene number in a biological pathway using almost complete saturation mutagenesis screening of haploid mouse cells. BMC Genomics. 2014;15(1016):1471–2164.
Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD. Construction of RFLP-based maps of foxtail millet, Setaria italica (L.) P. Beauv. Theor Appl Genet. 1998;96(1):31–6.
Weigel D. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiol. 2012;158(1):2–22. Epub 2011/12/08.
Yuan YW, Sagawa JM, Di Stilio VS, Bradshaw Jr HD. Bulk segregant analysis of an induced floral mutant identifies a MIXTA-like R2R3 MYB controlling nectar guide formation in Mimulus lewisii. Genetics. 2013;194(2):523–8. Epub 2013/04/09.

# Chapter 19
# Transposon Tagging in *Setaria viridis*

**Kazuhiro Kikuchi, Christine Shyu, and Thomas P. Brutnell**

**Abstract** Transposon tagging populations are important resources for characterizing gene function in model plant systems. In *Setaria viridis*, a transposon tagging population utilizing the *miniature Ping* (*mPing*) element was generated by introducing a construct containing *mPing* and the transposase *Ping* into accession A10.1. Transformants were screened for the presence of the transgene and *mPing* insertions were identified using nested inverse-PCR coupled with Illumina sequencing. Here we describe step-by-step protocols for generating the *Setaria viridis mPing* population, introduce methods for genotyping *mPing* insertion alleles, and discuss future implementations of the system.

**Keywords** mPing • Setaria • Transponson tagging • Inverse PCR

## 19.1 Introduction on Transposon Tagging and *mPing*

Transposon tagging is a powerful method to study gene function and has been utilized extensively in plant genetics. Transposons can be used in both forward and reverse genetic screens to generate mutations in known genes (Brutnell 2002; Izawa et al. 1997; Gierl and Saedler 1992). In general, transposon tagging exploits transposable elements to create novel genetic variation throughout the genome. In forward genetics, genes of interest can be cloned using sequences within the transposons as molecular tags and in reverse genetic screens, novel alleles of the gene of interest can be generated and studied. Populations are often maintained as independent lines containing a few to hundreds of single gene insertions that can serve as a rich resource for gene structure/function studies.

To exploit transposon tagging in *Setaria viridis*, we established a population of *Setaria viridis* (A10.1) plants harboring the rice *miniature Ping* (*mPing*) element. *mPing* is an active miniature inverted repeat transposable element (MITE) first

Kazuhiro Kikuchi and Christine Shyu contributed equally to the manuscript.

K. Kikuchi • C. Shyu • T.P. Brutnell, Ph.D. (✉)
Donald Danforth Plant Science Center, 975 North Warson Rd., St. Louis, MO 63132, USA
e-mail: tbrutnell@danforthcenter.org



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_19

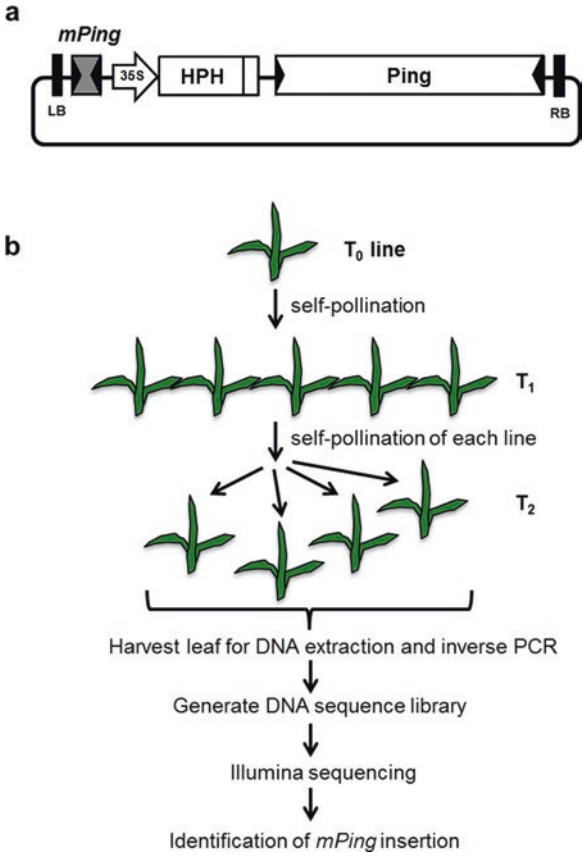


**Fig. 19.1** Overview of the development of *mPing* transposon-tagging populations. (**a**) Transgenic construct design. The construct was made using pBIN19 as the backbone. (**b**) Schematic overview for screening and identifying *mPing* insertion lines

identified in rice (Jiang et al. 2003; Kikuchi et al. 2003; Nakazaki et al. 2003). It is a nonautonomous element that transposes at high frequency and is activated by the autonomous elements *Ping* and *Pong*. *mPing* is 430 base pairs in length and preferentially transposes into promoter regions (Naito et al. 2009). Because of its small size and high transposition rate, *mPing* has been introduced into several species including soybean (Hancock et al. 2011), *Arabidopsis thaliana* (Yang et al. 2007), and yeast (Hancock et al. 2010) to serve as a mutagen.

To build an *mPing* population in *Setaria viridis*, a construct containing *mPing*, *Ping*, and a hygromycin-resistant cassette was transformed into accession A10.1 (See Fig. 19.1). Transformants were screened for the presence of the transgene, and nested inverse polymerase chain reaction (NiPCR) coupled with high-throughput sequencing was used to determine the *mPing* locations. In total, 41 independent $T_1$ lines containing *mPing* at multiple sites were generated as a resource for reverse genetics in *Setaria viridis*. Here we provide detailed protocols to develop a *Setaria viridis mPing* population and discusses future implementations of the system.

## 19.2 Methods and Protocol

### *19.2.1 Constructs and Generation of Transposon-Tagging Population*

To introduce *mPing* into *Setaria viridis*, carry out the following steps:

1. Build a construct containing the 430 bp *mPing* element and a full length *Ping* element with the target site duplication and terminal inverted regions from rice cv. Nipponbare into a pBIN19 plasmid (Bevan 1984) backbone carrying a *35S:: HYGROMYCIN PHOSPHOTRANSFERASE (HPT)* cassette (Fig. 19.1a).
2. Transform this construct into *Setaria viridis* using the transformation protocol described in Chap. 18 (Van Eck et al.).
3. Self-pollinate $T_0$ lines that are positive for the transgene and harvest seeds for further screening.
4. Plant individual $T_1$ seeds and collect $T_2$ seeds.
5. Plant 12 $T_2$ seeds from each of the $T_1$ plants and harvest leaf tissue from each $T_2$ individual to identify the location of the *mPing* insertion (Fig. 19.1b).

### *19.2.2 DNA Extraction*

1. Harvest three small pieces of young leaf tissue (about $3 \times 3$ mm$^2$ each) from each $T_2$ plant and place into individual wells containing three metal beads in a 96-well plate. Harvest samples from 12 $T_2$ seedlings from each independent $T_1$ line.
2. Add 200 μL KAZU buffer in each well and incubate in a 37 °C shaker at 300 rpm for 30 min to isolate DNA. Note: KAZU buffer is a newly developed DNA extraction buffer that significantly reduces time and cost of traditional DNA extraction methods and has been successfully used for DNA extraction of multiple grass species including *Setaria viridis* (KAZU buffer—Kerafast cat. #EDD003).
3. In a new plate, prepare a 1/8-in. whole punch of filter paper (Whatman) in each well. Add 20 μL of the solution onto the filter paper to bind DNA.
4. Use 100 μL of TE to wash the DNA. Repeat this washing step once.
5. Place DNA samples in a 65 °C dry bath for 10 min to evaporate excess solution from the filter paper.

### *19.2.3 Nested Inverse PCR*

To amplify regions in the genome where *mPing* is inserted, use nested inverse polymerase chain reaction (NiPCR) on DNA isolated from segregating $T_2$ individuals. iPCR is a variant of PCR that is used to amplify sequences that flank a region of known sequence and nested iPCR includes an additional PCR step to add barcodes

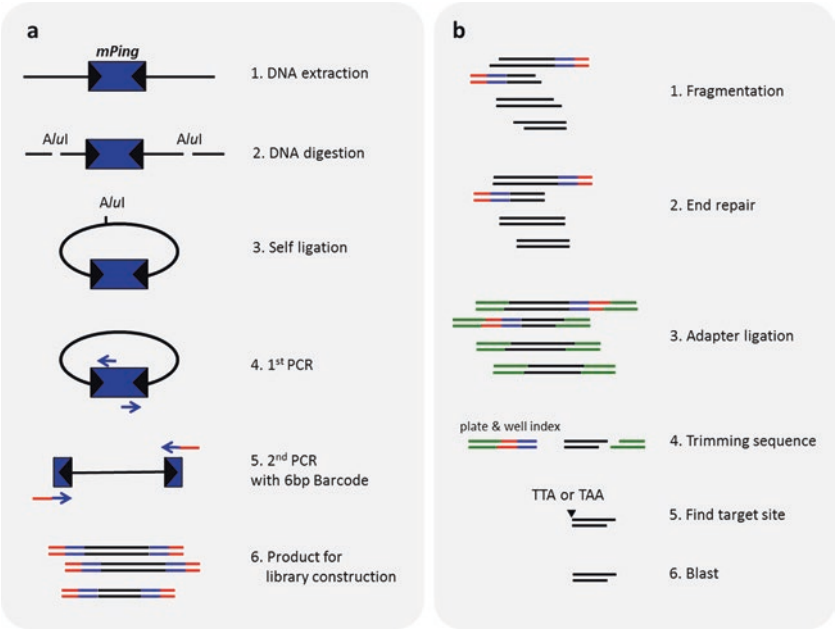


**Fig. 19.2** Schematic overview of nested inverse polymerase chain reaction (NiPCR), library construction, and bioinformatic analysis of the *Setaria viridis mPing* population. (**a**) NiPCR pipeline. *Blue arrows* in step 4 indicate primers that target subterminal regions of *mPing*. *Blue and red arrows* in step 5 indicate Terminal Inverted Repeat (TIR) primers that include 6 bp barcode sequences. (**b**) Library construction and bioinformatic analysis pipeline

onto the PCR amplicons (Ochman et al. 1988) (Fig. 19.2a). Following is the detailed protocol of NiPCR that includes DNA digestion, self-ligation, and two rounds of PCR amplification.

(a) DNA digestion

1. Add DNA (bound to dried filter paper; see previous section) to 1.4 mL matrix tubes (Thermo Scientific #4140) together with the restriction enzyme cocktail. For one reaction, use 3 μL NEBuffer, 0.5 μL AluI restriction enzyme (10,000 units/mL) and 26.5 μL nuclease-free water. Mix the sample well by pipetting.
2. Incubate the reaction at 37 °C for 2 h, then place matrix plate in incubator at 65 °C for 10 min to terminate the reaction.
3. Centrifuge the matrix plate for 30 s and continue immediately to the self-ligation step.

(b) Self-ligation

1. Prepare the master mix for the self-ligation reaction as follows (for a 100 sample reaction): 200 μL of T4 Ligase Buffer, 50 μL of T4 DNA Ligase and 750 μL of nuclease-free water.

2. Aliquot 10 μL of the master mix solution into each well. Then transfer 10 μL of restriction enzyme digested DNA solution to the new plate that contains the master mix. Make sure to mix the digested DNA sample thoroughly by pipetting up and down before transferring to the tubes with the master mix.
3. Centrifuge the plate for 5 s, then place in a thermocycler for an overnight incubation at 16 °C. If not continuing directly with the iPCR reaction, store the self-ligated solution at −20 °C.

(c) First PCR (iPCR I)

1. Prepare PCR reactions in a new 96-well plate containing the following reagents (for 100 samples): 25 μL GoTaq Polymerase, 500 μL 5× PCR GoTaq Buffer, 50 μL dNTP Mix, 250 μL Primer Duplex Mix (contains primers that target the subterminal region of *mPing*; Fig. 19.2a), and 1175 μL nuclease-free water.
2. Aliquot 20 μL of the master mix solution for each reaction and mix with 5 μL of self-ligated DNA into the corresponding well.
3. Mix the samples well, centrifuge, and amplify with the following cycles:

   (i) 95 °C 5 min
   (ii) 95 °C 30 s
   (iii) 50 °C 1 min
   (iv) 72 °C 1 min
   (v) Go to step (b) and repeat 29 times
   (vi) 72 °C 7 min
   (vii) 10 °C Infinite hold

4. PCR products can be stored at −20 °C.

(d) Second PCR (iPCR II)

1. Set-up a new 96-sample PCR Plate with a master mix containing the following reagents (for 100 samples): 50 μL GoTaq Polymerase (Promega #M300), 1000 μL PCR GoTaq Buffer, 100 μL dNTP Mix, 3350 μL nuclease-free water.
2. Mix 45 μL of the master mix solution with 2.5 μL of PCR products from the first PCR reaction.
3. Transfer 2.5 μL of Terminal Inverted Repeat (TIR) primer with unique barcode sequences (Table 19.1) into the PCR tubes. It is important to track each primer–sample combination for future informatics analysis.
4. Mix reaction well, centrifuge samples, and place into thermocycler using the following conditions:

   (i) 95 °C 5 min
   (ii) 95 °C 30 s
   (iii) 60 °C 1 min
   (iv) 72 °C 1 min
   (v) Go to step (ii) and repeat 29 times
   (vi) 72 °C 7 min
   (vii) 10 °C Infinite hold

**Table 19.1** Terminal inverted repeat (TIR) primer with unique barcode sequences used in second PCR

| Forward primer | Sequence 5′–3′ | Reverse primer | Sequence 5′–3′ |
|---|---|---|---|
| PING 5′TIR A1 | ATCATCACCCCCATTGTGACTGGCC | PING 3′TIR A1 | ATCATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR B1 | TACATCACCCCCATTGTGACTGGCC | PING 3′TIR B1 | TACATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR C1 | CACATCACCCCCATTGTGACTGGCC | PING 3′TIR C1 | CACATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR D1 | GACATCACCCCCATTGTGACTGGCC | PING 3′TIR D1 | GACATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR E1 | ATCATGACCCCCATTGTGACTGGCC | PING 3′TIR E1 | ATCATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F1 | TACATGACCCCCATTGTGACTGGCC | PING 3′TIR F1 | TACATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G1 | CACATGACCCCCATTGTGACTGGCC | PING 3′TIR G1 | CACATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H1 | GACATGACCCCCATTGTGACTGGCC | PING 3′TIR H1 | GACATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A2 | ATCGTCACCCCCATTGTGACTGGCC | PING 3′TIR A2 | ATCGTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR B2 | TACGTCACCCCCATTGTGACTGGCC | PING 3′TIR B2 | TACGTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR C2 | CACGTCACCCCCATTGTGACTGGCC | PING 3′TIR C2 | CACGTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR D2 | GACGTCACCCCCATTGTGACTGGCC | PING 3′TIR D2 | GACGTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR E2 | ATCGTGACCCCCATTGTGACTGGCC | PING 3′TIR E2 | ATCGTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F2 | TACGTGACCCCCATTGTGACTGGCC | PING 3′TIR F2 | TACGTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G2 | CACGTGACCCCCATTGTGACTGGCC | PING 3′TIR G2 | CACGTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H2 | GACGTGACCCCCATTGTGACTGGCC | PING 3′TIR H2 | GACGTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A3 | ATGATCACCCCCATTGTGACTGGCC | PING 3′TIR A3 | ATGATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR B3 | TAGATCACCCCCATTGTGACTGGCC | PING 3′TIR B3 | TAGATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR C3 | CAGATCACCCCCATTGTGACTGGCC | PING 3′TIR C3 | CAGATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR D3 | GAGATCACCCCCATTGTGACTGGCC | PING 3′TIR D3 | GAGATCACTAGCCATTGTGACTGGCC |
| PING 5′TIR E3 | ATGATGACCCCCATTGTGACTGGCC | PING 3′TIR E3 | ATGATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F3 | TAGATGACCCCCATTGTGACTGGCC | PING 3′TIR F3 | TAGATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G3 | CAGATGACCCCCATTGTGACTGGCC | PING 3′TIR G3 | CAGATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H3 | GAGATGACCCCCATTGTGACTGGCC | PING 3′TIR H3 | GAGATGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A4 | ATGCTCACCCCCATTGTGACTGGCC | PING 3′TIR A4 | ATGCTCACTAGCCATTGTGACTGGCC |

| Forward primer | Sequence 5′–3′ | Reverse primer | Sequence 5′–3′ |
|---|---|---|---|
| PING 5′TIR B4 | TAGCTCACCCCCATTGTGACTGGCC | PING 3′TIR B4 | TAGCTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR C4 | CAGCTCACCCCCATTGTGACTGGCC | PING 3′TIR C4 | CAGCTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR D4 | GAGCTCACCCCCATTGTGACTGGCC | PING 3′TIR D4 | GAGCTCACTAGCCATTGTGACTGGCC |
| PING 5′TIR E4 | ATGCTGACCCCCATTGTGACTGGCC | PING 3′TIR E4 | ATGCTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F4 | TAGCTGACCCCCATTGTGACTGGCC | PING 3′TIR F4 | TAGCTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G4 | CAGCTGACCCCCATTGTGACTGGCC | PING 3′TIR G4 | CAGCTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H4 | GAGCTGACCCCCATTGTGACTGGCC | PING 3′TIR H4 | GAGCTGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A5 | ACATCTACCCCCATTGTGACTGGCC | PING 3′TIR A5 | ACATCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B5 | TCATCTACCCCCATTGTGACTGGCC | PING 3′TIR B5 | TCATCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C5 | CTATCTACCCCCATTGTGACTGGCC | PING 3′TIR C5 | CTATCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D5 | GTATCTACCCCCATTGTGACTGGCC | PING 3′TIR D5 | GTATCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E5 | ACATCGACCCCCATTGTGACTGGCC | PING 3′TIR E5 | ACATCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F5 | TCATCGACCCCCATTGTGACTGGCC | PING 3′TIR F5 | TCATCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G5 | CTATCGACCCCCATTGTGACTGGCC | PING 3′TIR G5 | CTATCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H5 | GTATCGACCCCCATTGTGACTGGCC | PING 3′TIR H5 | GTATCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A6 | ACAGCTACCCCCATTGTGACTGGCC | PING 3′TIR A6 | ACAGCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B6 | TCAGCTACCCCCATTGTGACTGGCC | PING 3′TIR B6 | TCAGCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C6 | CTAGCTACCCCCATTGTGACTGGCC | PING 3′TIR C6 | CTAGCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D6 | GTAGCTACCCCCATTGTGACTGGCC | PING 3′TIR D6 | GTAGCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E6 | ACAGCGACCCCCATTGTGACTGGCC | PING 3′TIR E6 | ACAGCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F6 | TCAGCGACCCCCATTGTGACTGGCC | PING 3′TIR F6 | TCAGCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G6 | CTAGCGACCCCCATTGTGACTGGCC | PING 3′TIR G6 | CTAGCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H6 | GTAGCGACCCCCATTGTGACTGGCC | PING 3′TIR H6 | GTAGCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A7 | ACGACTACCCCCATTGTGACTGGCC | PING 3′TIR A7 | ACGACTACTAGCCATTGTGACTGGCC |

(continued)

**Table 19.1** (continued)

| Forward primer | Sequence 5′–3′ | Reverse primer | Sequence 5′–3′ |
|---|---|---|---|
| PING 5′TIR B7 | TCGACTACCCCCATTGTGACTGGCC | PING 3′TIR B7 | TCGACTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C7 | CTGACTACCCCCATTGTGACTGGCC | PING 3′TIR C7 | CTGACTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D7 | GTGACTACCCCCATTGTGACTGGCC | PING 3′TIR D7 | GTGACTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E7 | ACGACGACCCCCATTGTGACTGGCC | PING 3′TIR E7 | ACGACGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F7 | TCGACGACCCCCATTGTGACTGGCC | PING 3′TIR F7 | TCGACGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G7 | CTGACGACCCCCATTGTGACTGGCC | PING 3′TIR G7 | CTGACGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H7 | GTGACGACCCCCATTGTGACTGGCC | PING 3′TIR H7 | GTGACGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A8 | ACGTCTACCCCCATTGTGACTGGCC | PING 3′TIR A8 | ACGTCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B8 | TCGTCTACCCCCATTGTGACTGGCC | PING 3′TIR B8 | TCGTCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C8 | CTGTCTACCCCCATTGTGACTGGCC | PING 3′TIR C8 | CTGTCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D8 | GTGTCTACCCCCATTGTGACTGGCC | PING 3′TIR D8 | GTGTCTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E8 | ACGTCGACCCCCATTGTGACTGGCC | PING 3′TIR E8 | ACGTCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR F8 | TCGTCGACCCCCATTGTGACTGGCC | PING 3′TIR F8 | TCGTCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR G8 | CTGTCGACCCCCATTGTGACTGGCC | PING 3′TIR G8 | CTGTCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR H8 | GTGTCGACCCCCATTGTGACTGGCC | PING 3′TIR H8 | GTGTCGACTAGCCATTGTGACTGGCC |
| PING 5′TIR A9 | AGATGTACCCCCATTGTGACTGGCC | PING 3′TIR A9 | AGATGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B9 | TGATGTACCCCCATTGTGACTGGCC | PING 3′TIR B9 | TGATGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C9 | CGATGTACCCCCATTGTGACTGGCC | PING 3′TIR C9 | CGATGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D9 | GCATGTACCCCCATTGTGACTGGCC | PING 3′TIR D9 | GCATGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E9 | AGATGCACCCCCATTGTGACTGGCC | PING 3′TIR E9 | AGATGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR F9 | TGATGCACCCCCATTGTGACTGGCC | PING 3′TIR F9 | TGATGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR G9 | CGATGCACCCCCATTGTGACTGGCC | PING 3′TIR G9 | CGATGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR H9 | GCATGCACCCCCATTGTGACTGGCC | PING 3′TIR H9 | GCATGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR A10 | AGACGTACCCCCATTGTGACTGGCC | PING 3′TIR A10 | AGACGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B10 | TGACGTACCCCCATTGTGACTGGCC | PING 3′TIR B10 | TGACGTACTAGCCATTGTGACTGGCC |

| Forward primer | Sequence 5′–3′ | Reverse primer | Sequence 5′–3′ |
|---|---|---|---|
| PING 5′TIR C10 | CGACGTACCCCCATTGTGACTGGCC | PING 3′TIR C10 | CGACGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D10 | GCACGTACCCCCATTGTGACTGGCC | PING 3′TIR D10 | GCACGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E10 | AGACGCACCCCCATTGTGACTGGCC | PING 3′TIR E10 | AGACGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR F10 | TGACGCACCCCCATTGTGACTGGCC | PING 3′TIR F10 | TGACGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR G10 | CGACGCACCCCCATTGTGACTGGCC | PING 3′TIR G10 | CGACGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR H10 | GCACGCACCCCCATTGTGACTGGCC | PING 3′TIR H10 | GCACGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR A11 | AGTAGTACCCCCATTGTGACTGGCC | PING 3′TIR A11 | AGTAGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B11 | TGTAGTACCCCCATTGTGACTGGCC | PING 3′TIR B11 | TGTAGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C11 | CGTAGTACCCCCATTGTGACTGGCC | PING 3′TIR C11 | CGTAGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D11 | GCTAGTACCCCCATTGTGACTGGCC | PING 3′TIR D11 | GCTAGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E11 | AGTAGCACCCCCATTGTGACTGGCC | PING 3′TIR E11 | AGTAGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR F11 | TGTAGCACCCCCATTGTGACTGGCC | PING 3′TIR F11 | TGTAGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR G11 | CGTAGCACCCCCATTGTGACTGGCC | PING 3′TIR G11 | CGTAGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR H11 | GCTAGCACCCCCATTGTGACTGGCC | PING 3′TIR H11 | GCTAGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR A12 | AGTCGTACCCCCATTGTGACTGGCC | PING 3′TIR A12 | AGTCGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR B12 | TGTCGTACCCCCATTGTGACTGGCC | PING 3′TIR B12 | TGTCGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR C12 | CGTCGTACCCCCATTGTGACTGGCC | PING 3′TIR C12 | CGTCGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR D12 | GCTCGTACCCCCATTGTGACTGGCC | PING 3′TIR D12 | GCTCGTACTAGCCATTGTGACTGGCC |
| PING 5′TIR E12 | AGTCGCACCCCCATTGTGACTGGCC | PING 3′TIR E12 | AGTCGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR F12 | TGTCGCACCCCCATTGTGACTGGCC | PING 3′TIR F12 | TGTCGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR G12 | CGTCGCACCCCCATTGTGACTGGCC | PING 3′TIR G12 | CGTCGCACTAGCCATTGTGACTGGCC |
| PING 5′TIR H12 | GCTCGCACCCCCATTGTGACTGGCC | PING 3′TIR H12 | GCTCGCACTAGCCATTGTGACTGGCC |

5. iPCR II products now have unique barcodes for each sample. The products of a single 96-sample plate can now be pooled into a single 1.2 mL Eppendorf tube and stored for future library construction and DNA sequencing.

### *19.2.4 Illumina Library Construction*

To precisely map *mPing* insertion sites, prepare DNA sequencing libraries for Illumina sequencing (Fig. 19.2b). The procedure used for Illumina library construction includes PCR product pooling and cleanup, DNA fragmentation, end repair, dA-tailing, universal adapter ligation, PCR amplification, and size selection. Detailed protocols of each step are described as follows. High-throughput sequencing of NiPCR amplified fragments will yield *mPing*-junction fragments.

(a) PCR product pooling and cleanup

1. Take 10 μL of each iPCR II product from each well into a single 1.5 mL Eppendorf tube (plate pool). Label the tube with the corresponding 96-Well Matrix plate number (i.e., plate 1, 2, etc.).
2. Purify PCR products using a PCR purification kit (Qiagen #28104) with an additional wash step using 750 μL 80 % ethanol. For a 960 μL pool, elute purified PCR products in 100 μL of nuclease-free water and label as Q-DNA.
3. Store the Q-DNA samples at −20 °C for future library construction and DNA sequencing.

(b) Heat DNA fragmentation

1. Aliquot 20 μL of Q-DNA sample into five new PCR tubes and label each one "i–v". Incubate samples at 95 °C in a thermocycler.
2. Place each of the five tubes in the thermocycler for different time increments, for example:

   (i) 0 min at 95 °C
   (ii) 15 min at 95 °C
   (iii) 20 min at 95 °C
   (iv) 30 min at 95 °C
   (v) 45 min at 95 °C

3. Label the tube to correspond to the amount of time at 95 °C.
4. Load 10 μL of each sample onto a 2 % agarose gel together with a 100 bp and 1 kb DNA ladder to check fragmentation of samples (Fig. 19.3).
5. Mix the remaining 10 μL samples from each tube (i–v).
6. Store pooled sample at −20 °C for end-repair and library construction.

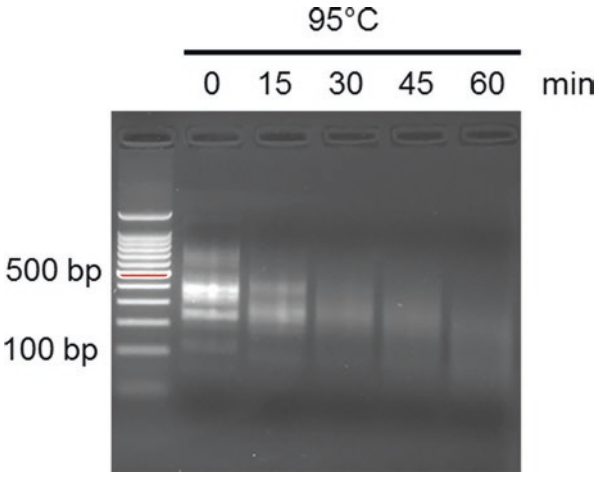


**Fig. 19.3** Gel electrophoresis of samples after the heat fragmentation step. Samples are then pooled for end repair

(c) End repair

1. Create a master mix that contains: 2 μL 10× NEBuffer No 2, 2.5 μL 1 mM dNTP mix, 0.5 μL (1U) Klenow (large fragment) enzyme, and 5.0 μL of nuclease-free water.
2. Add 10 μL of pooled sample from the heat DNA fragmentation step and mix well by pipetting.
3. Incubate samples at 37 °C for 30 min in a thermocycler followed by 10 min at 70 °C to terminate enzyme activity.

(d) dA-tailing

1. To the 20 μL sample, add the following: 3 μL of 10× NEBuffer No 2, 1 μL of 1 mM dATP, 0.5 μL of Klenow (3′–5′ exo) enzyme, and 25.5 μL of nuclease-free water. Mix thoroughly by pipetting.
2. Place the 50 μL sample at 37 °C for 30 min followed by 10 min at 70 °C to terminate enzyme activity in the thermocycler.

(e) Universal adapter ligation

1. Aliquot 10 μL of dA-tailed DNA into a new PCR tube and add the following: 1 μL of Illumina adapter (5 mM; Table 19.2), 2 μL 10× ligation buffer, 1 μL of T4 DNA Ligase, and 6 μL of nuclease-free water.
2. Incubate the mixture overnight at 16 °C in the thermocycler.

(f) PCR amplification

1. Aliquot 2 μL of self-ligated DNA (100 ng) into new PCR tubes. To each of the samples, add a mixture of the following: 5 μL of 5× Fidelity Buffer, 1 μL of 10 mM dNTP, 0.75 μL Index Primer (10 mM), 0.75 μL Universal Primer (10 mM), 0.5 μL KAPA Hi Fi Hot Start DNA Polymerase, and 15 μL nuclease-free water.
2. Place the 25 μL samples in thermocycler under the following conditions:

   (i) 95 °C 2 min

**Table 19.2** Primer sequences used in Illumina library construction

| Adapter and PCR primers | Sequence 5′–3′ |
|---|---|
| Universal adapter 1 | A*A*TGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCT TCCGAT*C*T |
| Universal adapter 2 | /5Phos/G*A*TCGGAAGAGCACACGTCTGAACTCCAGTC*A*C |
| Universal primer | AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC CGATCT |
| Index primer 1 | CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGT |
| Index primer2 | CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGT |
| Index primer3 | CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGT |
| Index primer4 | CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGT |
| Index primer5 | CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGT |
| Index primer6 | CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGT |
| Index primer7 | CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGT |
| Index primer8 | CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGT |
| Index primer9 | CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGT |
| Index primer11 | CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGT |
| Index primer12 | CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGT |
| Index primer13 | CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGACGTGT |
| Index primer14 | CAAGCAGAAGACGGCATACGAGATGGAACTGTGACTGGAGTTCAGACGTGT |
| Index primer15 | CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGACGTGT |
| Index primer16 | CAAGCAGAAGACGGCATACGAGATGGACGGGTGACTGGAGTTCAGACGTGT |
| Index primer17 | CAAGCAGAAGACGGCATACGAGATCTCTACGTGACTGGAGTTCAGACGTGT |

*Phosphorothioated DNA bases

(ii) 98 °C 20 s
(iii) 65 °C 30 s
(iv) 72 °C 20 s
(v) Go to step (ii) and repeat 15 times.
(vi) 72 °C 5 min
(vii) 12 °C Infinite hold

(g) Size selection
Carry out the following bead-based separation protocol to size-fractionate DNA fragments.

1. Prepare XP beads (Beckman Coulter, Inc cat #A63880) by vortexing thoroughly at a high speed and keep on ice while not in use.
2. To discard larger DNA fragments, use a 1:1 ratio of PCR product to XP beads by pipetting 25 μL of XP beads into the 25 μL of PCR product sample and mix thoroughly.
3. Incubate the mixture at room temperature for 6 min.
4. Place tubes on a magnetic strand and allow metal beads to be pulled completely out of the solution.
5. Transfer this supernatant (50 μL) into a new PCR tube. Discard the beads.
6. Bind smaller DNA fragments to beads by adding 50 μL of new XP beads to 50 μL of supernatant.
7. Incubate mixture at room temperature for 6 min.
8. Place tubes on the magnetic stand and allow metal beads to be pulled completely out of the solution.
9. Discard the supernatant carefully without losing any beads and wash twice with 75 % ethanol.
10. Let the beads dry in open air or using a small fan. Cautiously handle beads so that they are not disturbed when drying.
11. Elute the beads in 24 μL of Qiagen EB and mix the beads by pipetting thoroughly.
12. Incubate samples at room temperature for 2 min.
13. Place the tube on the magnetic stand and allow metal beads to be pulled out of the solution.
14. Transfer the supernatant into a new PCR tube and discard the beads. The supernatant is the final library and can be loaded on a 2 % agarose gel for quality control.
15. Samples can be stored at −20 °C before submitting for sequencing.

### *19.2.5 Illumina Sequencing and Analysis*

1. Sequence depth will depend on the complexity of the pool size. [Assuming 100 nt sequencing runs × 4 fragments × 96 PCR product × 100 × coverage = 3, 840,000 reads.]
2. To identify *mPing* insertion sites, assign samples to a plate using the adaptor/index sequence, and individual samples identified using the 6 bp barcode sequence. Only consider reads that carry both *mPing* and flanking sequences for *mPing* placement.
3. To reduce false positives, check for junction fragments that have the sequence TNA (*mPing*'s degenerate target site) immediately flanking the 3′ or 5′ TIR.

4. Align these reads to the genome of *Setaria viridis* (Phytozome v11) and map reads using the GSNAP program (Wu and Nacu 2010; Wu and Watanabe 2005).
5. Filter candidate insertions generated by this process to select high quality candidates for experimental validation. Consider putative insertion sites supported by less than 50 reads as low quality candidates and remove them from list. Merge remaining locations if predicted insertions sites are less than 50 base pairs apart from each other.
6. Locate the insertion genomic position isolate information on the annotated function of the nearest gene (if any) to the insertion site based on the *Setaria viridis* genomic sequence. In our experiment, 70 insertions were identified via this approach.

### *19.2.6 Confirm Genotyping and Characterization of Lines*

Use a PCR to verify *mPing* insertions in each individual line. Use primers targeting 150 bp upstream and downstream of the insertion site to check for the presence of *mPing*. A PCR product size shift of 430 base pairs is expected when *mPing* is inserted (Fig. 19.4). This PCR validation can be used to both verify the correct mapping of the *mPing* insertion as well as determine the zygosity of the insertion (one band—homozygous, two bands—hemizygous; Fig. 19.4). Alternatively, use a TIR primer (5′-CCATTGTGACTGGCC-3′) specific for both ends of *mPing* coupled with a gene-specific primer to verify insertion sites. PCR validation is best performed using DNA extracted from $T_2$ plants, then confirmed in $T_3$ families in order to determine if the insertion events were somatic or germinal events.

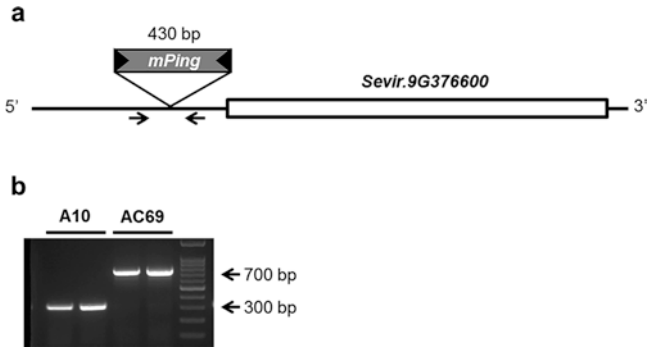


**Fig. 19.4** Example of *mPing* insertion in promoter region of gene of interest. (**a**) Cartoon of *mPing* line AC69 with *mPing* insertion in promoter region of *Sevir.9G376600*. (**b**) Genotyping PCR of *mPing* insertion in A10.1 and line AC69. Primer locations are indicated by *arrows* in (**a**). The results suggest that AC69 is homozygous for *mPing* in the promoter region of *Sevir.9G376600*

Because *mPing* is prone to insert in promoter regions, it is important to confirm alterations in gene expression. Quantitative-PCR is recommended to determine expression levels of the target gene before performing phenotypic analyses. Multiple backcrosses to A10.1 are recommended to minimize background effects.

## 19.3 Future Implementations

The *mPing* population not only broadens the tool set for reverse genetics in *Setaria viridis*, but also creates new possibilities for gene characterization given the preference for *mPing* to insert into promoter regions. Rapid development of genome editing techniques also opens new avenues for reverse genetic tools that can leverage *mPing* populations. Example future implementations are discussed as follows:

### *19.3.1 Engineering CRISPR-Cas9 Systems to Target mPing and Generate Additional Alleles*

Clustered, regularly interspaced, short palindromic repeat (CRISPR) technology is a rapidly developing technology for genome editing and engineering. It is composed of a guide RNA (gRNA) and a RNA-guided nuclease (such as Cas9) that targets genomic sequences complementing the gRNA for single- or double-stranded breaks [reviewed in Sander and Joung (2014)]. The precision of a CRISPR RNA-guided DNA–protein complex enables additional utility including transcriptional activation (CRISPRa), transcriptional repression (CRISPRi), or fluorescent tagging [reviewed in La Russa and Qi (2015)]. In transcriptional applications, the nuclease domains of the Cas9 protein are mutated, creating a nuclease-deactivated Cas9 (dCas9). dCas9 is fused to transcriptional activation domains such as VP64 (Konermann et al. 2013; Gilbert et al. 2013; Maeder et al. 2013; Perez-Pinera et al. 2013), transcriptional repression domains such as the Krüppel-associated box (KRAB) domain of Kox1 (Margolin et al. 1994), or fluorescent proteins (Chen et al. 2013) to achieve diverse functions upon targeting specific genomic regions. This opens new avenues to targeted transcriptional manipulation, which can be broadly applied to medical as well as agricultural engineering.

We envision applications for both CRISPRa and CRISPRi coupled to the *Setaria mPing* population to create novel variation for gene functional characterizations. As *mPing* preferentially inserts into promoter regions, constructs containing dCas-VP64 and gRNA sequences targeting *mPing* can be introduced into *mPing* lines to generate transgenic lines that have increased expression of genes flanking *mPing* insertions. Alternatively, dCas9 can be fused to the Ethylene-responsive element binding factor-associated amphiphilic repression (EAR) motif, which is an important transcriptional repression domain broadly present in hormonal and developmental

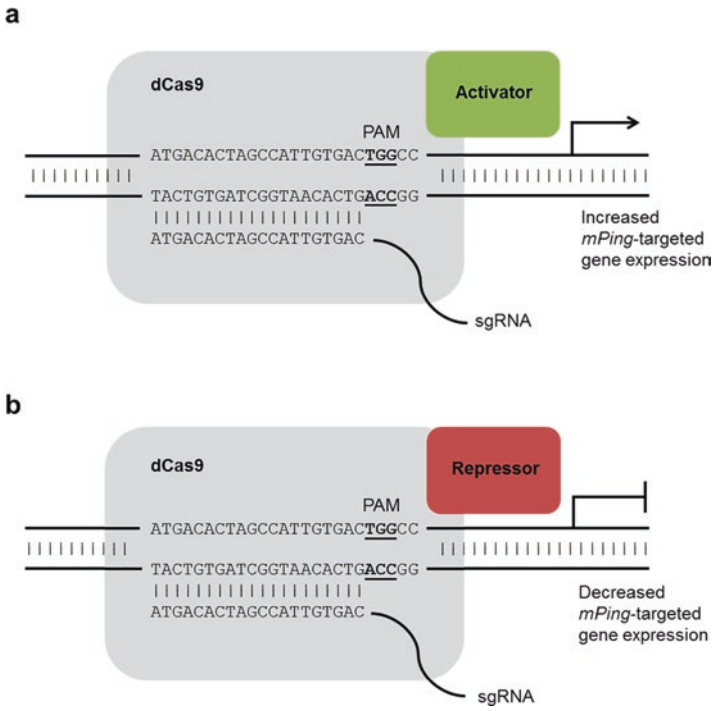


**Fig. 19.5** Utilizing CRISPRa (**a**) and CRISPRi (**b**) to activate or repress *mPing*-targeted gene expression, respectively. (**a**) CRISPRa. dCas9 will be fused with an activation domain such as VP64 to activate downstream transcription. A gRNA targeting the 3′ end of TIR in the *mPing* element will guide the dCas9-activator complex to *mPing* regions. (**b**) CRISPRi. Similar to CRISPRa, a dCas9 fusion protein with transcriptional repressor domains will be guided to *mPing* regions through an *mPing*-specific gRNA for targeted gene repression

signaling pathways in plants (Kagale and Rozwadowski 2011), to create alleles that have reduced transcriptional expression. A combination of dCas9-VP64 and dCas-EAR lines along with the original *mPing* line can thus create an array of transgenic lines that have different expression levels of the targeted gene of interest (Fig. 19.5). Because gene expression levels in these systems are all regulated by trans-acting dCas9 proteins, genomic DNA sequences of the target gene are identical between lines, creating a more homogenous background for phenotypic comparison. Moreover, although CRISPR systems theoretically can target a very broad range of genomic sequences, gRNA efficiencies often vary (Hsu et al. 2013; Bortesi and Fischer 2015). Optimizing gRNA sequences that target *mPing* thus provides the advantage of high-efficiency targeting of any *mPing* line, creating a library of standard dCas9 constructs that can be used broadly in *mPing* populations. The dCas9 cassettes could be introduced by transformation or through simple crosses once stock lines have been generated. Taken together, this system combining *mPing* with dCas9 constructs would provide important building blocks for generating a broad range of transgenic lines that are crucial for studying gene function.

### 19.3.2 Engineering mPing for Stress-Inducible Gene Tagging

Given that *mPing* preferentially inserts into promoter regions, direct engineering of the *mPing* sequence also allows manipulation of downstream gene expression. *mPing* is composed of 430 base pairs that includes 15 base pair terminal inverted repeats at both ends of the sequence (Kikuchi et al. 2003). Cis-element scanning of the *mPing* sequence identified multiple stress-responsive regulatory motifs, and it has been hypothesized that *mPing* plays a role in providing binding sites for transcription factors to induce gene expression upon stress (Naito et al. 2009). However, in addition to stress-responsive regulatory motifs, there is also a wide array of hormonal and developmental regulatory motifs, potentially blurring the response. Thus, engineering *mPing* to contain repetitive stress-responsive cis-elements such as the abscisic acid-responsive element (ABRE) or dehydration-responsive element (DRE) (Yamaguchi-Shinozaki and Shinozaki 1994, 2005) will allow stronger and more specific stress-inducible gene expression of the *mPing* target. Building populations with these ABRE- or DRE-enriched *mPing* elements will expand the toolkit for *Setaria* genetics and potentially identify novel genes involved in stress responses.

In addition to stress-responsive cis-elements, hormonal-responsive or cell-type specific regulatory domains can also be engineered in *mPing* to accommodate different biological questions. These resources can allow identification and characterization of genes involved in specific regulatory pathways while minimizing the chance of having lethal lines when critical genes are altered at the whole-plant level. Though engineering *mPing* to contain constitutive promoters can be very appealing, most constitutive promoters utilized in grass systems such as the maize ubiquitin promoter 1 (Christensen and Quail 1996) and rice actin promoter (McElroy et al. 1990) are well over 400 base pairs, exceeding the length for endogenous *mPing* to actively transpose. However, detailed functional characterization of *mPing* sequences and an understanding of the *mPing* transposition mechanism are at a very early stage. Exploring the maximum length of sequences that can be engineered into *mPing* without affecting *mPing* excision and transposition activity, as well as characterization of the minimal sequence for constitutive expression in grass systems will allow more sophisticated manipulations of the *mPing* system.

### 19.3.3 Expanding mPing Populations with Existing mPing Lines

Although only 41 independent *mPing* lines are available in the current Setaria *mPing* population, additional lines can be generated through introducing autonomous elements through crossing or transformation. *mPing* has the ability to attain copy numbers as high as 1000 elements in the rice genome (Naito et al. 2006). Though detailed mechanisms of *mPing* transposition are yet to be explored, it is understood that the autonomous

transposases *Ping* and *Pong* play critical roles in *mPing* excision (Yang et al. 2007). The Setaria *mPing* population was initially generated with a full length *Ping* element (Fig. 19.1). However, *Ping* activity was lost after three generations, resulting in stable *mPing* lines for characterization (Kazihiro Kikuchi, unpublished). Reintroducing *Ping* into the *mPing* population has the potential to reactivate *mPing* transposition, thus expanding the population and creating additional *mPing* insertion lines that target a broader percentage of genes. Once the desired number of independent lines are obtained, *Ping* elements can then be crossed out to stabilize insertions. This system can be broadly applied to *mPing* populations with endogenous *mPing* sequences as well as engineered *mPing* sequences as mentioned in the previous section.

## 19.4 Summary and Outlook

This protocol describes the construction of an *mPing* population in *Setaria viridis* (A10.1), providing a first demonstration of *mPing* movement in a panicoid grass and laying a foundation for reverse genetics. The system utilizes the MITE *mPing*, which preferentially inserts into promoter regions. A construct containing *mPing*, an autonomous *Ping*, and a hygromycin-resistant cassette was transformed into *Setaria viridis* (A10.1) and populations screened for *mPing* transposition. NiPCR coupled with high-throughput sequencing was performed to determine *mPing* insertion sites. Finally, a population of 41 independent lines was developed that contains multiple *mPing* insertions. This database largely utilizes molecular tools developed for *Setaria viridis* and serves as a valuable addition to genetic resources. This system provides the basis for a wide array of implementations utilizing characterized stress regulatory elements, CRISPRa and CRISPRi technology, as well as further manipulation of the *mPing* system itself. Resources generated from transposon tagging in *Setaria* will complement rapidly developing technologies such as RNA-interference and CRISPR-Cas technology, greatly broadening the toolsets available in *Setaria viridis* and strengthening *Setaria viridis* as a model for panicoid grasses.

**Acknowledgements** The authors would like to thank James Schnable for mapping reads to the Setaria genome, and undergraduate interns Alexandra Chapovskaya, Dante Espero, Amanda Orosco, and Daniel Perez for assisting with genotyping. This work was supported by National Science Foundation (IOS-1027542) to K.K. and T.P.B., and United States Department of Agriculture-National Institute of Food and Agriculture (2014-67012-22269) to C.S.

## References

Bevan M. Binary Agrobacterium vectors for plant transformation. Nucleic Acids Res. 1984;12(22): 8711–21.

Bortesi L, Fischer R. The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol Adv. 2015;33(1):41–52.

Brutnell TP. Transposon tagging in maize. Funct Integr Genomics. 2002;2(1–2):4–12.
Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013;155(7):1479–91.
Christensen AH, Quail PH. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res. 1996;5(3):213–8.
Gierl A, Saedler H. Plant-transposable elements and gene tagging. Plant Mol Biol. 1992;19(1):39–49.
Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013;154(2):442–51.
Hancock CN, Zhang F, Wessler SR. Transposition of the Tourist-MITE mPing in yeast: an assay that retains key features of catalysis by the class 2 PIF/Harbinger superfamily. Mob DNA. 2010;1(1):5.
Hancock CN, Zhang F, Floyd K, Richardson AO, LaFayette P, Tucker D, et al. The rice miniature inverted repeat transposable element mPing is an effective insertional mutagen in soybean. Plant Physiol. 2011;157(2):552–62.
Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31(9):827–32.
Izawa T, Ohnishi T, Nakano T, Ishida N, Enoki H, Hashimoto H, et al. Transposon tagging in rice. Plant Mol Biol. 1997;35(1–2):219–29.
Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, et al. An active DNA transposon family in rice. Nature. 2003;421(6919):163–7.
Kagale S, Rozwadowski K. EAR motif-mediated transcriptional repression in plants: an underlying mechanism for epigenetic regulation of gene expression. Epigenetics. 2011;6(2):141–6.
Kikuchi K, Terauchi K, Wada M, Hirano H-Y. The plant MITE mPing is mobilized in anther culture. Nature. 2003;421(6919):167–70.
Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, et al. Optical control of mammalian endogenous transcription and epigenetic states. Nature. 2013;500(7463):472–6.
La Russa MF, Qi LS. The new state of the art: CRISPR for gene activation and repression. Mol Cell Biol. 2015;35(22):3800–9.
Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. Nat Methods. 2013;10(10):977–9.
Margolin JF, Friedman JR, Meyer WK, Vissing H, Thiesen HJ, Rauscher III FJ. Kruppel-associated boxes are potent transcriptional repression domains. Proc Natl Acad Sci U S A. 1994;91(10):4509–13.
McElroy D, Zhang W, Cao J, Wu R. Isolation of an efficient actin promoter for use in rice transformation. Plant Cell. 1990;2(2):163–71.
Naito K, Cho E, Yang G, Campbell MA, Yano K, Okumoto Y, et al. Dramatic amplification of a rice transposable element during recent domestication. Proc Natl Acad Sci U S A. 2006;103(47):17620–5.
Naito K, Zhang F, Tsukiyama T, Saito H, Hancock CN, Richardson AO, et al. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature. 2009;461(7267):1130–4.
Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H, et al. Mobilization of a transposon in the rice genome. Nature. 2003;421(6919):170–2.
Ochman H, Gerber AS, Hartl DL. Genetic applications of an inverse polymerase chain reaction. Genetics. 1988;120(3):621–3.
Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods. 2013;10(10):973–6.
Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32(4):347–55.
Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics. 2010;26(7):873–81.
Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics. 2005;21(9):1859–75.

Yamaguchi-Shinozaki K, Shinozaki K. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. Plant Cell. 1994;6(2):251–64.

Yamaguchi-Shinozaki K, Shinozaki K. Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. Trends Plant Sci. 2005;10(2):88–94.

Yang G, Zhang F, Hancock CN, Wessler SR. Transposition of the rice miniature inverted repeat transposable element mPing in Arabidopsis thaliana. Proc Natl Acad Sci U S A. 2007;104(26):10962–7.

# Chapter 20
# *Agrobacterium tumefaciens*-Mediated Transformation of *Setaria viridis*

**Joyce Van Eck, Kerry Swartwood, Kaitlin Pidgeon, and Kimberly Maxson-Stein**

**Abstract** Gene transfer methodology, often referred to as transformation, is an important component of a model system research platform. Availability of transformation methods markedly expand the application of a model. We chose to develop an *Agrobacterium tumefaciens*-mediated gene transfer method for *Setaria viridis* A10.1. Regenerable callus was recovered from mature seeds without seed coats that were disinfected and cultured on a Murashige and Skoog-based medium supplemented with 40 g/L maltose, 2 mg/L 2,4-dichlorophenoxyacetic acid, 0.5 mg/L kinetin, and 4 g/L Gelzan. The gelling agents Gelzan and Phytagel were found to be critical for recovery of a high quality callus as compared to agar that resulted in a gelatinous, brown, non-regenerable callus. For transformation, the callus was infected with the *A. tumefaciens* strain AGL1 that contained binary vectors with the hygromycin phosphotransferase selectable marker gene, which confers resistance to the antibiotic hygromycin. The transformation efficiency, which is defined as the percent of infected callus that gives rise to at least one independent transgenic line ranged from 0.3 to 15 % depending upon the vector backbone used to design constructs and the gene of interest that was either overexpressed or had a knockdown of expression. Transgenic lines were first verified by PCR, then positive plants were moved forward for copy number determination by either Southern or TaqMan® analysis. On average, 42 % of the transgenic lines contained one copy of the introduced transgene. Availability of a transformation methodology has contributed to the adoption of *S. viridis* as a model species by researchers worldwide.

**Keywords** Copy number • Gene transfer • Green foxtail • Monocot transformation • *Setaria*

J. Van Eck (✉) • K. Swartwood • K. Pidgeon
Boyce Thompson Institute, 533 Tower Road, Ithaca, NY 14853, USA
e-mail: jv27@cornell.edu; kec37@cornell.edu; kep67@cornell.edu

K. Maxson-Stein
Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132, USA
e-mail: KMaxson-Stein@danforthcenter.org



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_20

## 20.1 Introduction

### *20.1.1 General Information on Gene Transfer Methodology*

Plant biotechnology had its beginnings when tissue culture methods for generation of plants from cells were successfully combined with the ability to transfer genes of interest into plant cells. This greatly advanced studies related to gene identification and function with model and crop species, which has led to an enhanced understanding of gene networks and mechanisms. Gene transfer into plant cells has been achieved through direct DNA uptake into protoplasts (Paszkowski et al. 1984), delivery of DNA by particle bombardment (Sanford et al. 1987), and by *Agrobacterium tumefaciens*-mediated transformation (Gelvin 2003). Each of these approaches has advantages and disadvantages that need to be considered when choosing a gene transfer method for a plant species of interest. The direct DNA uptake approach requires methods for plant regeneration from protoplasts, which can be difficult for many plant species. For particle bombardment, a gene gun must be available, which can be cost prohibitive, and for *Agrobacterium*-mediated transformation the plant tissue needs to be amenable to infection by the bacterium and also not be adversely affected.

We chose to use *Agrobacterium tumefaciens* for development of a transformation methodology for *S. viridis* because we routinely use it for transformation of other plant species; therefore, we have the experience and all the necessary resources. *A. tumefaciens* is a soil borne plant pathogenic bacterium that engineers the cells it infects to produce metabolites needed for its survival. Galls are produced at the sites of infection and eventually as the galls enlarge the plant's phloem and xylem become compressed, which leads to the death of the plant.

Approximately 30 years ago, scientists used an altered form of an *A. tumefaciens* strain to produce transgenic lines of petunia, tobacco, and tomato in the laboratory (Horsch et al. 1985). In the years since that report, there have been an impressive number of reports of transformation of both dicots and monocots; however, initially, efficient transformation of monocots proved to be more problematic (Sood et al. 2011). This was due in part to a lack of efficient plant regeneration methodologies at that time and, in addition, monocots are not natural hosts of *A. tumefaciens*. Over the years, these barriers were overcome and successful *Agrobacterium*-mediated transformations of monocots have become more commonplace. There have been reports of various target tissues used for *Agrobacterium* infection of monocots that include immature embryos (Ishida et al. 2007) and regenerable callus derived from immature embryos, immature inflorescences, and mature seeds (Burris et al. 2009; Lee et al. 2004; Song et al. 2011).

### *20.1.2 Background on the Process to Develop Methods for* Setaria viridis

In this chapter, we describe development of methods for plant regeneration from mature seed-derived callus of *Setaria viridis* and *Agrobacterium tumefaciens*-mediated transformation using that callus as the target tissue. An earlier version of this method

was described in a report regarding the future of *S. viridis* as a model species for studies of $C_4$ photosynthesis (Brutnell et al. 2010). Since publication of that early report, we have made changes to the methods (Van Eck and Swartwood 2015).

The first step in development of the *S. viridis* transformation methodology was to establish a tissue culture approach that would be highly efficient for plant regeneration. Having an efficient regeneration system would improve the chances of recovering transgenic plant lines. We reviewed the literature for reports on tissue culture methods for *S. viridis*, but the only reports we found were for its close relative, *S. italica* (Rao et al. 1988; Rout et al. 1998; Samantaray et al. 1999). In these reports, methods were outlined for production of a plant regenerable callus from mature seeds. Mature seed-derived callus has also been used for *Agrobacterium*-mediated transformation of other monocots such as rice (Ozawa 2009), switchgrass (Li and Qu 2011), and perennial ryegrass (Patel et al. 2013). In addition to the *S. italica* reports regarding seed-derived callus, there are also reports on callus derived from immature inflorescences (Reddy and Vaidyanath 1990; Vishnoi and Kothari 1996; Xu et al. 1984).

After reviewing the literature, we chose to initiate experiments with *S. viridis* according to the methods outlined in the reports for production of a regenerable *S. italica* mature-seed derived callus. There are advantages to using mature seeds as a starting material for callus when compared to immature inflorescences. Seeds can be easily stored, are readily available, and in our experience with other plant species are more easily disinfected than immature inflorescences. We can generate a large supply of *S. viridis* seeds that can in turn produce a larger amount of callus than is possible with immature inflorescences. In addition, there is less risk from the introduction of insects such as thrips and spider mites into a tissue culture facility from mature seeds than immature inflorescences freshly harvested from plants grown in a growth chamber, greenhouse, or the field. Although the inflorescence material would undergo disinfection with decontaminant solutions of ethanol and bleach, the eggs of insects are sometimes impervious to these chemicals, and crevices throughout the inflorescences may provide protection for these insects if the decontaminant solutions do not reach deep enough.

## 20.2 Establishment of Methods for Production of Regenerable Callus from Mature Seeds

### *20.2.1 Methods to Break Dormancy*

Based on the advantages of mature seed for production of regenerable callus, we chose to follow published reports on *S. italica* although we made several modifications in order to recover a good quality, highly regenerable callus (Fig. 20.1). Although there are advantages to using mature seeds, we found there were problems with seed dormancy for the *S. viridis* A10.1 germplasm if the seeds were younger than 3 months old. We found that the most effective method to break seed dormancy was mechanical removal of the seed coats, which did not damage the embryos and


**Fig. 20.1** Plant regeneration from *Setaria viridis* A10.1 mature seed-derived callus. (**a**) Regenerable callus derived from mature seeds without seed coats that were cultured on callus induction medium. (**b**) Production of shoots from callus that was cultured on plant regeneration medium. (**c**) Rooted plant recovered - a shoot removed from callus on regeneration medium was transferred to rooting medium

resulted in approximately 100 % germination (Van Eck and Swartwood 2015). Other methods that have been found to be effective are treatment with liquid smoke (Sebastian et al. 2014) and storage of seeds in moist sphagnum moss at 4 °C for 3–6 weeks (H. Jiang, personal communication).

### *20.2.2 Callus Induction*

We found the most efficient approach to generate a highly regenerable callus was a two-step method, firstly culturing mature seeds on a callus induction medium (CIM) and then transferring the resultant callus to plant regeneration medium (PRM). Before culturing the seeds, the seed coats were removed according to Van Eck and Swartwood (2015) then disinfected for 3 min in a solution of sterile deionized water, 10 % (v/v) bleach and 0.1 % (v/v) Tween 20 followed by three rinses with sterile

water. The CIM was a Murashige and Skoog (MS)-based medium (Murashige and Skoog 1962) that contained the growth regulators kinetin (0.5 mg/L) and 2,4-dichorophenoxyacetic acid (2 mg/L) (Van Eck and Swartwood 2015). In order to recover a high quality, regenerable callus, we found it was critical to incorporate a gellan gum such as Gelzan (4 g/L) or Phytagel (4 g/L) to solidify the medium because when agar was used the majority of the callus produced was brown, gelatinous, and non-regenerable.

### *20.2.3 Physical and Environmental Parameters*

In addition to the effects of medium composition on callus development, we also identified several physical and environmental conditions during the culture process that affected the development of callus from the mature seeds. We found that placing the seeds on CIM with the embryos facing upwards had a positive effect on callus development when compared with placing the embryos down in contact with the medium. It was also important to limit the number of seeds cultured on CIM to 15 per Petri plate (100 mm × 20 mm) because we saw a reduction in the amount of callus produced as the number of seeds per plate was increased to 25.

We found that maintenance at 24 °C resulted in the highest callus quality, which is unlike some monocots such as rice where maintenance at 28 °C is common. When we compared light versus dark conditions at 24 °C, we observed a negative effect of light on the quality of the callus. Also important for recovery of a high quality callus was a lack of condensation in the Petri plates for both callus initiation and maintenance.

## 20.3 Plant Regeneration

We determined that an MS-based medium designated PRM containing kinetin (0.5 mg/L) and Phytoblend (7 g/L) as the gelling agent resulted in multiple shoots per callus from 100% of the seed-derived calli (Fig. 20.1). Originally, we used either Gelzan or Phytagel to solidify the regeneration medium because this was part of the CIM, but occasionally we observed hyperhydricity or water-soaked appearance (Kevers et al. 2004) of the regenerated plants.

However, when we substituted Phytablend this hyperhydricity was eliminated. There have been other reports of the occurrence of hyperhydricity depending on the plant species and gelling agent used to solidify medium (Ivanova and Van Staden 2011). When callus was transferred to PRM, the cultures were moved to light conditions (16-h photoperiod, 57–65 μE $m^{-2}$ $s^{-1}$) and we continued to maintain the cultures at 24 °C.

## 20.4 Rooting Regenerated Shoots In Vitro

Regenerated shoots were transferred to a medium designated RM (regeneration medium) that contained half-strength MS salts, full-strength MS vitamins, and Phytoblend (7 g/L) (Van Eck and Swartwood 2015). Both Magenta™ GA7 containers and test tubes worked well as culture vessels for the rooting stage. Shoots rooted within a week of transfer to rooting medium and were maintained under light conditions (16-h photoperiod, 57–65 μE $m^{-2}$ $s^{-1}$) at 24 °C (Fig. 20.1).

## 20.5 Transfer of In Vitro Plants to Soil

When the in vitro plants were well rooted, they were transferred to a commercially available soil mix and then moved into a growth chamber. The culture medium was carefully washed from the roots with tepid water before transfer to moistened Metro Mix 360 Plus (Hummert International, Earth City, MO) in 10.2 $cm^2$ horticultural plant containers (smaller size containers can also be used). Each plant was covered with a transparent, plastic, reused beverage bottle that had the bottom removed (Fig. 20.2). After 1–2 days the caps were removed from the bottles and the bottles were removed 2–3 days later. This format was followed to provide a gradual acclimation from in vitro to in vivo conditions. The plants were

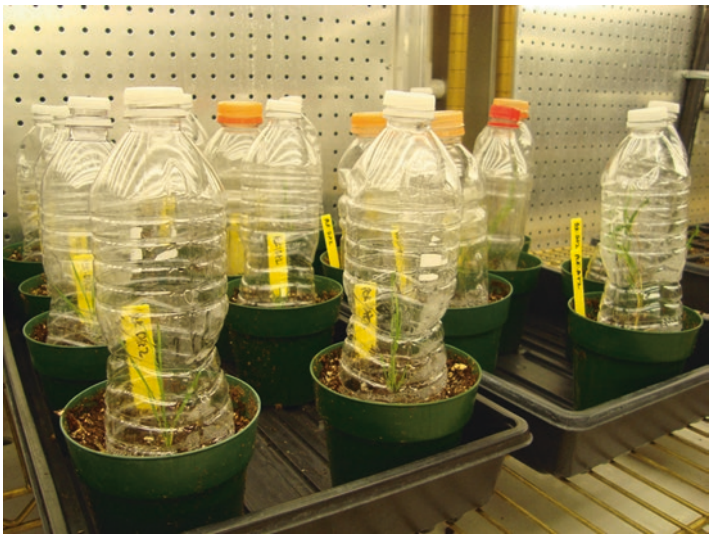

**Fig. 20.2** Regenerated plants of *Setaria viridis* A10.1 after transfer from in vitro culture conditions to a growth chamber. Plants were transferred to a soil mix and covered with reused beverage bottles for a gradual acclimation to growth chamber conditions. The bottle caps were removed after 1–2 days and the bottles were removed 2–3 days later

maintained in a growth chamber at 23 °C with a 16-h photoperiod at 210 μE $m^{-2}$ $s^{-1}$. Plants began to flower 1–3 weeks after transfer to the soil mix, with seeds ready to harvest approximately 5 weeks later.

## 20.6 Gene Transfer Method Development

### *20.6.1 Agrobacterium tumefaciens Strain and Vectors*

After we developed highly efficient and repeatable plant regeneration methods, we began investigating *Agrobacterium tumefaciens*-mediated transformation (Brutnell et al. 2010; Van Eck and Swartwood 2015). To develop the methods, we used the *A. tumefaciens* strain AGL1 that contained either one of two binary vectors. The first vector we used was pOL001 (Vogel and Hill 2008) that contains both the hygromycin phosphotransferase (*hpt*) plant selectable marker gene and the beta-glucuronidase (GUS) (Jefferson et al. 1987) reporter gene. The other binary vector used was pWBVec8, which is the parent vector of pOL001 and contains the *hpt* gene alone (Wang et al. 1998). The *hpt* gene confers resistance to the antibiotic hygromycin, therefore, to select for transgenic events hygromycin was incorporated into CIM, PRM, and RM. After testing the efficacy for selection of various hygromycin concentrations in each medium, we found the following concentrations to be optimum: CIM 40 mg/L; PRM 15 mg/L; RM 20 mg/L.

Beyond method development, we investigated transformation with other vectors and constructs that contained genes of interest. We found the backbone of the vector to be an important consideration for successful *S. viridis* transformation. Additional vectors besides pOL001 and pWBVec8 that worked well were the pMDC and pANIC series, which includes vectors for both overexpression and knockdown of expression (Curtis and Grossniklaus 2003; Mann et al. 2012).

Three to four days prior to infection, AGL1 containing a construct of interest was streaked from glycerol stocks onto plates of MG/L medium (Per liter: 5 g tryptone, 2.5 g yeast extract, 5 g NaCl, 5 g mannitol, 0.1 g $MgSO_4$, 0.25 g $K_2HPO_4$, 1.2 g glutamic acid, 15 g sucrose, pH to 7.2, 15 g Bacto Agar) containing the appropriate antibiotics. The plates were maintained at 28 °C for 48 h or until colonies appeared. Colonies were used to inoculate a liquid culture of LB medium (Per liter: 10 g tryptone, 5 g yeast extract, 10 g NaCl) containing the appropriate antibiotics and grown to an $OD_{600}$ of 0.6 in a shaking incubator (28 °C, 125 rpm). The culture was placed in a centrifuge at 8000 rpm for 10 min and the resultant pellet was resuspended in liquid CIM containing 33 μg/mL acetosyringone and 1 mg/mL Synperonic PE/F68 (Sigma-Aldrich Co., St. Louis, MO).

### 20.6.2 Infection and Cocultivation of Seed-Derived Callus with Agrobacterium

One week prior to infection with *Agrobacterium*, 6–7-week-old callus was divided into 2–3 mm pieces and placed onto freshly prepared CIM. For infection, 50 calli were transferred to a 15 mL centrifuge tube, the prepared infection solution (see Sect. 20.6.1) was added, and then allowed to incubate at room temperature for 5 min.

Cocultivation occurred in plates of CIM that had 7 cm filter paper placed on the surface of the medium. The cultures were maintained in the dark at 22 °C for 3 days. We investigated longer cocultivation periods, but they did not result in a significant increase in the recovery of transgenic lines. Following the cocultivation period, the calli were transferred to CIM Selective (CIMS) medium that contained 40 mg/L hygromycin and were maintained in the dark at 24 °C.

### 20.6.3 Recovery of Putative Transgenic Shoots and Rooting

For recovery of putative transgenic shoots, we investigated various times of maintenance on CIMS and found that 16 days were optimum before transfer to PRM Selective (PRMS) containing 15 mg/L hygromycin. When calli were transferred to PRMS, they were moved to light conditions as described in Sect. 20.3. Once shoots developed and were 0.8 cm tall they were removed from the callus and transferred to selective RM medium (RMS) containing 20 mg/L hygromycin. Only one shoot was selected from each infected callus to ensure that we were selecting truly independent transgenic lines and not sister clones that may have arisen from the infection of the same cell. Rooted shoots were analyzed by PCR for the transgene before moving forward for additional analysis. These primary transgenics are designated T0 and subsequent generations from self-pollinations of *S. viridis* are referred to as T1, T2, etc. The time required from infection of calli to recovery of T1 seeds is approximately 4 months.

To determine the transformation efficiency, we divided the number of PCR positive transgenic lines by the number of infected calli. The transformation efficiency ranged from 0.3 to 15 % depending upon the vector backbone and the gene of interest that was either overexpressed or had expression knocked down via RNA interference (RNAi) constructs.

## 20.7 Molecular Analyses of *Setaria viridis* Transgenic Lines

For molecular evaluation of *S. viridis* transgenic lines, we chose to utilize three different methods of analysis depending upon the level of characterization necessary. PCR was done as a first level of analysis on rooted in vitro plants to verify that

putative transgenics did indeed contain the introduced transgene. PCR positive plants were moved forward for further assessment. To determine that PCR positive plants recovered from the same *Agrobacterium*-infected callus were indeed independent transgenic lines, we used Southern blot analysis to survey differences in integration patterns of the transgene. Results from Southern blot analysis were also used to evaluate copy number of the introduced transgene, however in addition, we developed an alternative method for copy number determination of *S. viridis* transgenic lines based on a quantitative real-time PCR assay called TaqMan®.

### *20.7.1 PCR Analysis*

DNA was isolated from leaves of in vitro rooted plants and used for PCR reactions according to the Extract-N-Amp™ plant PCR kit (Sigma-Aldrich). We used a duplex set of PCR primers for each sample tested. One primer set was designed to detect the *hpt* selectable marker gene. The forward primer sequence was 5′-AGG CTC TCG ATG AGC TGA TGC TTT-3′ and the reverse primer sequence was 5′-AGC TGC ATC ATC GAA ATT GCC GTC-3′, which resulted in a product size of 335 bp. Along with the *hpt* primer set, we also included a primer set specific for *S. viridis* DNA as an internal control to make certain the DNA was of a quality and quantity to result in PCR product. Therefore, if we observed a product with the *Setaria*-specific primers, but did not observe a product with the *hpt* primers, we had confidence that a putative transgenic line was actually negative for the introduced transgene, and it was not a problem with the DNA preparation. The forward primer sequence for the *S. viridis*-specific primers was 5′-CAG CAA GCC GCC TAT ATG GAG-3′, and the reverse sequence was 5′-TCG TCT CAG GAG TGG CCA AGT C-3′ with a product size of 540 bp.

### *20.7.2 Southern Blot and TaqMan® Analysis*

Determining the copy number of an introduced transgene is important because high-copy number has been shown to affect expression not only of the introduced gene but also endogenous genes (Hobbs et al. 1993; Stam et al. 1997). It is also essential for government regulatory approval because release of low-copy number events is a critical consideration. The goal of our work was to recover transgenic lines with three copies or less of the introduced transgene.

We followed a CTAB method for DNA isolation and found that it was necessary to use very young leaves because isolation from older leaves resulted in preparations that had a high concentration of polysaccharides. These polysaccharides caused difficulty in quantification of the DNA and inhibition of restriction enzyme digest for Southern analysis. For T0 plants, we waited at least 1 week after the plants were transferred from in vitro conditions to a soil mix before harvesting leaves. For T1 plants, we harvested leaves from 10- to 12-day-old plants.

The DNA extraction buffer contained 100 mM Tris-Cl (pH 8.0), 20 mM EDTA (pH 8.0), 1.4 M NaCl, 2 % (w/v) CTAB, and 1 % (w/v) PVP 40,000 (polyvinyl pyrolidone). The solvent for phase separation contained Phenol/Chloroform/Isoamyl Alcohol (25:24:1 Mixture, pH 6.7/8.0) (Thermo Fisher Scientific, Waltham, MA) and was supplemented with additional Tris-Cl buffer to achieve pH 8.0. The resultant DNA was dissolved in 50 µL 10 mM Tris buffer, pH 8.5, and stored at −20 °C for short-term storage and −80 °C for long-term storage (3 or more months).

For Southern blot analysis, we used a nonradioactive chemiluminescence method which has been successfully used on plant DNA (McCabe et al. 1997) that relies on the use of a digoxigenin (DIG) label to detect nucleic acid (dUTP) that is incorporated into the hybridization probe (McCreery 1997). Compared to the radioactive detection method, the DIG method has several advantages: it is safer due to the lack of handling radioactive materials, labeled probes can be stored up to 1 year, space does not need to be allocated for a restricted area for the work, there is no hazardous waste disposal, and the DIG method requires shorter exposure times.

We used 8 µg of total DNA per digest for the Southern blot analysis of each transgenic line. A variety of restriction enzymes with 6 bp recognition sites were successfully used in combination with probes for the *hpt*, GUS, the green fluorescent protein (GFP) reporter gene, and GUS linker sequence. Figure 20.3 contains an image of a Southern blot of *Hind*III-digested *S. viridis* DNA, hybridized with a DIG-labeled *hpt* probe. Copy number for each transgenic line can be determined based on the number of bands visible in each lane. All of these transgenic lines resulted from transformations with constructs designed with the same vector backbone.

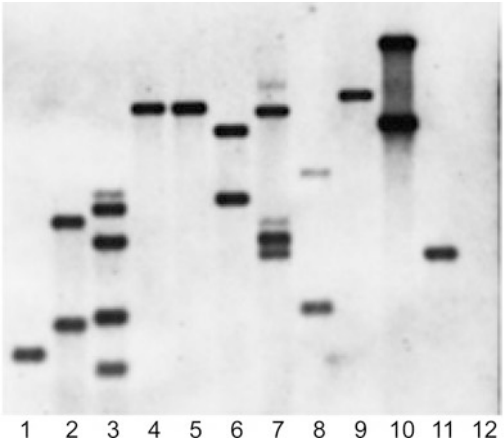


**Fig. 20.3** Southern blot results from the $T_0$ generation of transgenic *S. viridis*. 8–10 µg of DNA was digested with *Hind*III and probed with approximately 500 bp of DIG-labeled *hpt* sequence. *Lanes 1–11* represent T0 lines with different copy numbers of the introduced transgene. *Lane 12* is wild-type accession A10.1. *Lanes 1, 4, 5, 9, and 11* show single-copy events. *Lanes 4 and 5* show clonal events evident from similar banding patterns in two individuals regenerated from the same callus

The introduced transgene ranged from 1 to 5 copies with the majority of the lines containing 1 or 2 copies. As indicated earlier, Southern blot analysis can also be used to verify that transgenic lines are indeed independent events based on the banding pattern, which indicates integration of the transgene. Lanes 4 and 5 contain DNA from two lines that were recovered from the same callus and because the banding pattern is similar it can be concluded that these lines are not independent transformation events but arose from a single infected cell.

The TaqMan® method for copy number determination is a quantitative PCR approach that requires a real-time PCR machine (Bubner and Baldwin 2004; Ingham et al. 2001). This method is less labor and time intensive than Southern blot analysis. Two advantages of the TaqMan® assay for determining copy number in *S. viridis* transgenic lines are that only 50 ng of genomic DNA are needed per reaction compared with 8 μg for Southerns, and the TaqMan® assay takes only 2 h to produce results as compared with 1 week for Southerns. However, if you do not have access to a real-time PCR machine, Southerns are a good alternative for determining copy number, as the quality of the results are similar.

The TaqMan® analysis was performed in a duplex PCR including primers and probes for both the *hpt* selectable marker gene and an internal reference. The internal reference was the *S. viridis* phosphoenolpyruvate carboxykinase (*pck*) gene reported to be single copy in the genome (Xu et al. 2013). In addition, TaqMan® probes were each labeled with a 5′ fluorophore and a 3′ quencher for selective detection and quantification of specific target amplicons. The following are sequences with labels we designed for *hpt* and *pck* (Life Technologies, Grand Island, NY):

| | | |
|---|---|---|
| Hpt F3 | 5′-ATACGAGGTCGCCAACATCT | |
| Hpt R3 | 5′-CGTCTGCTGCTCCATACAAG | |
| Hpt Probe3 | 5′-CCAACCACGGCCTCCAGAAGA | 5′FAM™ 3′QSY® |
| SvPCK F3 | 5′-ACTGTGGACGGCATAAAGG | |
| SvPCK R3 | 5′-CGTACTTGGTCGGATGGAG | 5′VIC® 3′TAMARA™ |
| SvPCK Probe3 | 5′-ACGTTCTCGGCCTGCTTCGG | |

Overall, copy number assessment by both TaqMan® and Southern analyses showed that the average number of single-copy transgenic lines recovered from transformation experiments with different vector backbones was 42 %.

Once we identified low-copy number T0 transgenic lines, they were advanced to the T1 generation. At this stage, we wanted to determine zygosity of the transgene because of the necessity to work with homozygous populations to characterize gene function. Established methods to determine zygosity rely on the identification of non-segregating T2 populations by either germination of T2 seed on selective medium or a genotyping PCR on the T2. As an alternative, we used the TaqMan® analysis to determine zygosity of single-copy and linked multiple-copy lines in the T1 generation (German et al. 2003). To verify the TaqMan® analysis of zygosity, we compared the TaqMan® analysis of five independent events in the T1 generation with a genotyping PCR for segregation in the corresponding T2 generation.

With the TaqMan® analysis, we were able to identify homozygous lines 100 % of the time. We found the TaqMan® method to be less labor and time intensive compared with established methods to determine zygosity. The ability to streamline the identification of homozygous lines early in the T1 generation is key for advancing material for downstream studies.

## 20.8 Summary

Adoption of *S. viridis* as a model has grown worldwide since the first report that described its advantages as a model system for studies of the cellular and molecular mechanisms of $C_4$ photosynthesis (Brutnell et al. 2010). As is evident in the reports throughout this book, the development of resources for *S. viridis* has also increased in parallel with the interest in this new monocot model. Information gathered from investigations that utilize *S. viridis* will contribute to a better understanding of the biology and biochemical components of $C_4$ photosynthesis and characteristics of interest for development of biofuels crops. What is learned from the studies of the genetic regulation of $C_4$ photosynthesis has the potential through genetic engineering to increase yield and confer durability to withstand extreme abiotic stresses in important $C_3$ photosynthesis species such as rice and wheat. Development of *S. viridis* transformation methodology will continue with a focus on making the process as high-throughput as possible, which in turn will decrease the time and associated costs of labor and supplies to identify genes of interest and their function in metabolic networks.

**Acknowledgements** The authors acknowledge Dr. Erica Unger-Wallace for providing the *Setaria viridis* specific PCR primer sequences and also Christina Azodi for the initial efforts on development of the TaqMan® methods. We also acknowledge our funding sources, the Association of Independent Plant Research Institutes (AIPI) and the Triad Foundation.

## References

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu XG, Kellogg E, Van Eck J. *Setaria viridis*: a model for C4 photosynthesis. Plant Cell. 2010;22:2537–44.

Bubner B, Baldwin IT. Use of real-time PCR for determining copy number and zygosity in transgenic plants. Plant Cell Rep. 2004;23:263–71.

Burris JN, Mann DGJ, Joyce BL, Stewart CN. An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). Bioenerg Res. 2009;2:267–74.

Curtis MD, Grossniklaus U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol. 2003;133:462–9.

Gelvin SB. *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev. 2003;67:16–37.

German MA, Kandel-Kfir M, Swarzberg D, Matsevitz T, Granot D. A rapid method for the analysis of zygosity in transgenic plants. Plant Sci. 2003;164:183–7.
Hobbs SLA, Warkentin TD, Delong CMO. Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol Biol. 1993;21:17–26.
Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. A simple and general-method for transferring genes into plants. Science. 1985;227:1229–31.
Ingham DJ, Beer S, Money S, Hansen G. Quantitative real-time PCR assay for determining transgene copy number in transformed plants. Biotechniques. 2001;31:132–40.
Ishida Y, Hiei Y, Komari T. *Agrobacterium*-mediated transformation of maize. Nat Protoc. 2007;2: 1614–21.
Ivanova M, Van Staden J. Influence of gelling agent and cytokinins on the control of hyperhydricity in *Aloe polyphylla*. Plant Cell Tiss Organ Cult. 2011;104:13–21.
Jefferson RA, Kavanagh TA, Bevan MW. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 1987;6:3901–7.
Kevers C, Franck T, Strasser RJ, Dommes J, Gaspar T. Hyperhydricity of micropropagated shoots: a typically stress-induced change of physiological state. Plant Cell Tiss Organ Cult. 2004;77: 181–91.
Lee SH, Lee DG, Woo HS, Lee BH. Development of transgenic tall fescue plants from mature seed-derived callus via *Agrobacterium*-mediated transformation. Asian Aust J Anim Sci. 2004;17:1390–4.
Li RY, Qu RD. High throughput *Agrobacterium*-mediated switchgrass transformation. Biomass Bioenerg. 2011;35:1046–54.
Mann DG, Lafayette PR, Abercrombie LL, King ZR, Mazarei M, Halter MC, Poovaiah CR, Baxter H, Shen H, Dixon RA, Parrott WA, Stewart Jr NC. Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. Plant Biotechnol J. 2012;10:226–36.
McCabe MS, Power JB, de Laat AMM, Davey MR. Detection of single-copy genes in DNA from transgenic plants by nonradioactive Southern blot analysis. Mol Biotechnol. 1997;7:79–84.
McCreery T. Digoxigenin labeling. Mol Biotechnol. 1997;7:121–4.
Murashige T, Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant. 1962;15:473–97.
Ozawa K. Establishment of a high efficiency *Agrobacterium*-mediated transformation system of rice (*Oryza sativa* L.). Plant Sci. 2009;176:522–7.
Paszkowski J, Shillito RD, Saul M, Mandak V, Hohn T, Hohn B, Potrykus I. Direct gene transfer to plants. EMBO J. 1984;3:2717–22.
Patel M, Dewey RE, Qu RD. Enhancing *Agrobacterium tumefaciens*-mediated transformation efficiency of perennial ryegrass and rice using heat and high maltose treatments during bacterial infection. Plant Cell Tiss Organ Cult. 2013;114:19–29.
Rao AM, Kishor PBK, Reddy LA, Vaidyanath K. Callus induction and high-frequency plant-regeneration in Italian millet (*Setaria italica*). Plant Cell Rep. 1988;7:557–9.
Reddy LA, Vaidyanath K. Callus formation and regeneration in two induced mutants of foxtail millet (*Setaria italica*). J Genet Breed. 1990;44:133–8.
Rout GR, Samantaray S, Das P. *In vitro* selection and characterization of Ni-tolerant callus lines of *Setaria italica* L. Acta Physiol Plant. 1998;20:269–75.
Samantaray S, Rout GR, Das P. *In vitro* selection and regeneration of zinc tolerant calli from *Setaria italica* L. Plant Sci. 1999;143:201–9.
Sanford JC, Klein TM, Wolf ED, Allen N. Delivery of substances into cells and tissues using a particle bombardment process. Particul Sci Technol. 1987;5:27–37.
Sebastian J, Wong MK, Tang E, Dinneny JR. Methods to promote gemination of dormant *Setaria viridis* seeds. PLoS ONE. 2014; 9:e95109.
Song G, Walworth A, Hancock JF. Factors influencing *Agrobacterium*-mediated transformation of switchgrass cultivars. Plant Cell Tiss Organ Cult. 2011;108:445–53.

Sood P, Bhattacharya A, Sood A. Problems and possibilities of monocot transformation. Biol Plant. 2011;55:1–15.

Stam M, Mol JNM, Kooter JM. The silence of genes in transgenic plants. Ann Bot. 1997;79:3–12.

Van Eck J, Swartwood K. *Setaria viridis*. In: Wang K, editor. *Agrobacterium* protocols, vol. 2. New York: Springer; 2015. p. 57–67.

Vishnoi RK, Kothari SL. Somatic embryogenesis and efficient plant regeneration in immature inflorescence culture of *Setaria italica* (L) Beauv. Cereal Res Commun. 1996;24:291–7.

Vogel J, Hill T. High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. Plant Cell Rep. 2008;27:471–8.

Wang MB, Li ZY, Matthews PR, Upadhyaya NM. Improved vectors for *Agrobacterium tumefaciens*-mediated transformation of monocot plants. Acta Hortic. 1998;461:401–7.

Xu ZH, Wang DY, Yang LJ, Wei ZM. Somatic embryogenesis and plant-regeneration in cultured immature inflorescences of *Setaria italica*. Plant Cell Rep. 1984;3:149–50.

Xu J, Li Y, Ma X, Ding J, Wang K, Wang S, Tian Y, Zhang H, Zhu X-G. Whole transcriptome analysis using next-generation sequencing of model species *Setaria viridis* to support C4 photosynthesis research. Plant Mol Biol. 2013;83:77–87.

# Chapter 21
# Spike-Dip Transformation Method of *Setaria viridis*

**Prasenjit Saha and Eduardo Blumwald**

**Abstract** Conventional *Agrobacterium*-mediated transformation of monocots involves labor-intensive and time-consuming in vitro tissue culture methodology to regenerate T0 plants. To overcome these difficulties, a simple *in planta Agrobacterium*-mediated genetic transformation method for the emerging monocot model *Setaria viridis* was developed. Initial standardization of transient and stable transformations was performed using *A. tumefaciens* strain AGL1 harboring the *β-glucuronidase* (*GUS*) reporter gene driven by the *cauliflower mosaic virus 35S* (*CaMV35S*) promoter to transform preanthesis developing spikes. The method was further optimized by using *A. tumefaciens* strain EHA105 carrying *β-glucuronidase plus (GUSplus)*, green fluorescent protein (*GFP*), and *Discosoma* sp. red fluorescent protein (*DsRed*) reporter genes driven by either *CaMV35S* or an intron-interrupted maize *ubiquitin* (*Ubi*) promoters to develop stable transgenic lines from *S. viridis.* Dipping of 5-day-old S3 spikes into *Agrobacterium* cultures containing *S. viridis* spike-dip medium supplemented with 0.025 % Silwet L-77 and 200 μM acetosyringone for 20 min produced stable transformants at the rate of 0.8 ± 0.1 %. Transgenic lines showed stable integration of transgenes into the genome, and inherited transgenes followed the Mendelian segregation pattern and were expressed in subsequent generations. This spike-dip method will facilitate high-throughput translational research in a monocot model.

**Keywords** *Agrobacterium*-mediated transformation • Emerging monocot model • *In planta* • *Setaria viridis* • Spike-dip

P. Saha • E. Blumwald (✉)
Department of Plant Sciences, University of California,
One Shields Avenue, Davis, CA 95616, USA
e-mail: psaha@ucdavis.edu; eblumwald@ucdavis.edu



A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3_21

## 21.1 Introduction

Monocot $C_4$ cereal grain and grasses belonging to the Panicoideae subfamily such as *Zea mays*, *Sorghum bicolor*, S*accharum officinarum*, *Pennisetum glaucum*, *Panicum virgatum*, etc., which represent major sources of global food and feedstock, are gaining increased research interest because of the rising demand of food and fuel security (Sage and Zhu 2011). *Setaria viridis* with its small size, simple growth requirements, short life cycle, and small genome not only provides a model for $C_4$ photosynthetic research (Brutnell et al. 2010), but is also ideal for functional genomics studies to investigate abiotic stress tolerance in Panicoideae crops (Saha et al. 2016).

Although a traditional method of *Agrobacterium*-mediated transformation, through the involvement of tissue culture for *S. viridis,* has been reported (Brutnell et al. 2010; Eck and Swartwood 2015; Martins et al. 2015b) (Chap. 20), the development of an alternative transformation procedure for this model plant that eliminates the tissue culture and regeneration phases would enable significant progress in testing gene functions.

Traditional method of *Agrobacterium*-mediated transformation of *S. viridis*, an emerging monocot model, involves labor-intensive and time-consuming in vitro tissue cultures where transformed undifferentiated callus tissue becomes organogenic to regenerate whole T0 plants (Brutnell et al. 2010; Eck and Swartwood 2015) (Chap. 20). The amenability of in vitro cultured tissues to *Agrobacterium*-mediated transformation is often genotype dependent, and frequently produce somaclonal variations or genetic chimeras with morphological abnormalities and reduced fertility, including significant epigenetic alterations (Wang and Wang 2012). Clough and Bent (1998) developed an alternative method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*, a dicot model species, using an *in planta* floral-dip that obviates the in vitro tissue culture phase. Subsequently, a similar *Agrobacterium*-mediated floral transformation approach was developed for *Camelina sativa* (Liu et al. 2012), *Linum usitatissimum* (Bastaki and Cullis 2014), *Raphanus sativus* (Curtis and Nam 2001), *Solanum lycopersicum* (Yasmeen et al. 2009), and the model legume *Medicago truncatula* (Trieu et al. 2000). An *in planta* floral-dip method was reported for *Zea mays* (Mu et al. 2012), a major monocot food crop with a large genome belonging to the Panicoideae (Schnable et al. 2009), a close relative of *S. viridis*. Unfortunately, such *in planta* floral-dip method is not available for *S. viridis* with a relatively small true diploid genome (~510 Mb) (Bennetzen et al. 2012). Recently, the feasibility of the floral-dip transformation of *S. viridis* was suggested (Martins et al. 2015a), but the method was not optimized, the transformation efficiency was not determined, and reproducibility of the method was not established.

An optimized alternative method for the *Agrobacterium*-mediated transformation of *S. viridis* using a spike-dip that avoids the traditional in vitro culture steps has been developed (Saha and Blumwald 2016). This method produced fertile transgenic T1 plants within 8–10 weeks and the Mendelian pattern of inherited transgenes expressed over generations. The protocol provides an *in planta* monocot transformation system and will be widely applicable for functional genomics studies.

## 21.2 *S. viridis* Genotypes and Growth Conditions for Spike-Dip Transformation

Five *S. viridis* genotypes, which included three tolerant [A10.1 (PI 669942/Ames 31045), 132 (PI Ames 28193), and 98HT-80 (PI 649320)] and two sensitive to water-deficit and high-temperature stresses [Dekker 1851 (PI 223677) and PI 408811 (UI 4833)] (Saha et al. 2016), were taken under investigation for spike-dip transformation (Table 21.1). Spike-dip transformation of three tolerant genotypes has been reported recently (Saha and Blumwald 2016), while transformation of two sensitive genotypes is under progress. Seeds of these genotypes were obtained from the Germplasm Resources Information Network (GRIN, http://www.ars-grin.gov/), United States Department of Agriculture (USDA) and germinated in trays (27.9 cm × 54.3 cm) (McConkey, Sumner, WA) containing moist agronomy mix (equal parts of redwood compost, sand, and peat moss). Trays were kept in an Isotemp incubator (Fisher Scientific, Pittsburgh, PA) to vernalize at 4 °C for 2 days and then moved at 28° ± 2 °C with 50 % relative humidity for 16 h day/8 h night photoperiod in the greenhouse. After 7 days postgermination (DPG), seedlings were transferred to pots (10.2 cm × 8.2 cm) (McConkey) containing moist agronomy mix and allowed to grow under greenhouse conditions until spike development (Fig. 21.1a). During this

**Table 21.1** Development and progress of spike-dip transformation using *Agrobacterium tumefaciens* strain EHA105 harboring the *35S::GUS* (pCAMBIA1201) reporter gene and genes conferring tolerance to water-deficit and high-temperature

| Accession | Genotype | Origin (country/area) | Transformation efficiency ± SD | Phenotype under stress[a] | References |
|---|---|---|---|---|---|
| Ames 28193 | 132 | Kazakhstan | 0.7 ± 0.0 | Tolerant | Saha et al. (2016) and Saha and Blumwald (2016) |
| PI 669942/ Ames 31045 | A10.1 | United States | 0.7 ± 0.2 | Tolerant | Saha et al. (2016) and Saha and Blumwald (2016) |
| PI 649320 | 98HT-80 | Mongolia | 0.8 ± 0.1 | Tolerant | Saha et al. (2016) and Saha and Blumwald (2016) |
| PI 223677 | Dekker1851 | Azerbaijan | ND | Sensitive | Saha et al. (2016) |
| PI 408811 | UI4833 | China | ND | Sensitive | Saha et al. (2016) |

[a]Phenotype of wild-type plants under water-deficit and high temperature stress conditions; *SD* standard deviation, *ND* not determined (transformation in progress)

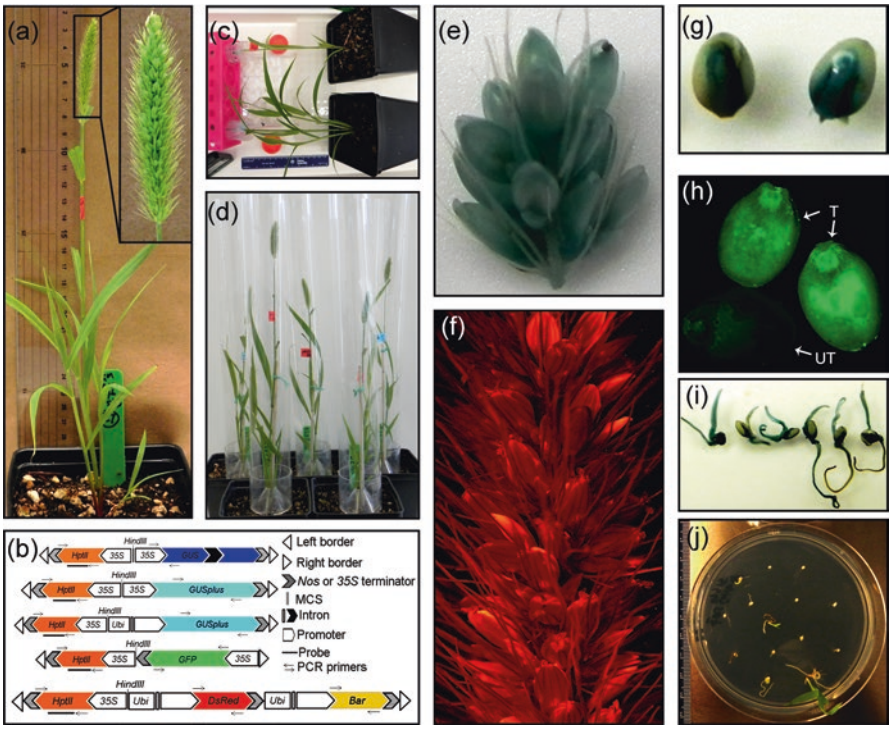


**Fig. 21.1** Step of Spike-dip transformation method of *Setaria viridis* A10.1. (**a**) A *S. viridis* plant with primary tiller bearing single inflorescence with an enlarged view of a spike at S3 stage prior to spike-dip transformation. (**b**) Schematic representation of linearized maps of the T-DNA cassettes of plant transformation vectors. Reporter [*β-glucuronidase (GUS)*, *β-glucuronidase plus (GUSplus)*, green fluorescent protein (*GFP*) and *Discosoma* sp. red fluorescent protein (*DsRed*) reporters)], and *hygromycin phosphotransferase* (*HptII*), bialaphos (*Bar*) selectable marker genes are driven by *CaMV35S* (*35S*) and maize *ubiquitin* (*Ubi*) promoters and terminated by *CaMV35S* or *nopaline synthase* (*NOS*) terminators. The position of gene-specific PCR primers, location of *Hind*III restriction enzyme within the multiple cloning site (MCS) or T-DNA cassette and the probe for Southern blot analysis are represented. (**c**) Top view of *S. viridis* plants during Spike-dip transformation. (**d**) Plants in arecones after 5 days of posttransformation inside growth room/chamber. (**e**) Histochemical GUS staining of portion of spike transformed with *35S::GUS* (pCAMBIA1201) reporter gene construct after 3 days postdip. (**f**) Transient expression of *DsRed* reporter gene from spike transformed with *Ubi::DsRed* (pGWB17-UbiDsRED-UbiBAR) after 5 days postdip. (**g**) Histochemical GUS staining of mature T1seeds transformed with *35S::GUS* (pCAMBIA1201) reporter gene. (**h**) Expression of *GFP* reporter gene in mature T1 seeds transformed with *35S::GFP* (pH7m24GW35Sp-GFP) construct (*T* transformed, *UT* untransformed). (**i**) Histochemical assay for *GUSplus* reporter gene activity in 3 days post germinated seedling transformed with *Ubi::GUSplus* (pH7m24GW pUbi-GUSplus) construct. (**j**) Selection of putative transformants at the T1 plant generation on plates containing half MS + hygromycin (30 mg/L)

time plants were watered every alternate day with deionized water and fertilized using a solution of 50 % N: phosphorus: potassium (20:10:20) and 50 % ammonium sulfate (total of 0.5 g of N) once in a week until maturity.

## 21.3 Suitable Spike Developmental Stage for *Agrobacterium*-Mediated Spike-Dip Transformation

The developmental stage of the flowers was shown to be crucial for successful floral transformation (Clough and Bent 1998; Desfeux et al. 2000; Trieu et al. 2000; Zale et al. 2009). High rates of transformation in *Arabidopsis* occurred when the plants had a maximum number of unopened floral bud clusters, and that the dipping of flowers later than 4 days before anthesis did not produce transformants (Clough and Bent 1998). The development of *S. viridis* accession A10.1 spikes from early preanthesis to late postanthesis can be categorized into seven stages (S1 to S7) based on the first visualization of their emergence among leaf sheath, auricle, and flag leaf (Saha and Blumwald 2016). The early S2 stage is suitable for transient transformation while S3 spikes of approximately 5.5 cm long during 6–8 days before anthesis are most amenable to *Agrobacterium* transformation (Fig. 21.1a). With the progression of spike development toward anthesis (S4 to S6), the transformation rates decreased gradually (Saha and Blumwald 2016). The preanthesis spikes of 6–7 cm long at its early or mid-uninucleate microspore stage which did not emerge from the sheath before 4–7 days of anthesis are the most susceptible target tissues amenable to *Agrobacterium* in wheat (Zale et al. 2009). High histochemical GUS staining at the bottom of each floret where the ovary is located indicated that the ovule is the probable target tissue amenable to T-DNA transfer upon *Agrobacterium* infection. Bechtold and Pelletier (1998), Desfeux et al. (2000), and (Ye et al. 1999) showed that the *Arabidopsis* ovule is the likely target for T-DNA transfer. Multiple *Agrobacterium* inoculations by three consecutive dipping of the same spikes caused bleaching of plant tissues with no seed setting as most spikes died (Saha and Blumwald 2016).

## 21.4 Appropriate Medium for Optimum Spike-Dip Transformation

The nutrient-rich *S. viridis* Spike Dip (*Sv*SD) medium containing 10.5 g/L $K_2HPO_4$, 4.5 g/L $KH_2PO_4$, 1 g/L $(NH4)_2SO_4$, 0.5 g/L NaCitrate, 4 g/L glycerol, 1 mM $MgSO_4$, 15 g/L abscorbic acid, 10 mM MES; pH to 5.8 (Lee and Yang 2006), with modification of sugar constituents to glucose (36 g/L) and sucrose (68.5 g/L) produced two to four-fold higher transformation rates when tested using *A. tumefaciens* strain AGL1 carrying *35S::GUS* reporter gene construct compared to other dipping media such as *S. viridis* Infiltration (*Sv*I) medium containing 50 mM MES, 2 mM $Na_3PO_4$, $12H_2O$, 0.025% (w/v) glucose; pH 5.8 (Brutnell et al. 2010); Murashige and Skoog (MS) full strength basal medium; pH 5.8 (Murashige and Skoog 1962); Sucrose (5% w/v) (Bent 2006); and MS full strength basal medium fortified with 5% (w/v) sucrose; pH 5.8 (Clough and Bent 1998). The combination of glucose and sucrose was also found to be effective for high transformation rates in conventional transformation methods of rice (Mohanty et al. 1999; Saha et al. 2006b). Glycerol and sucrose were found to be critical components of the *Sv*SD medium and elimination of either component from the *Sv*SD medium

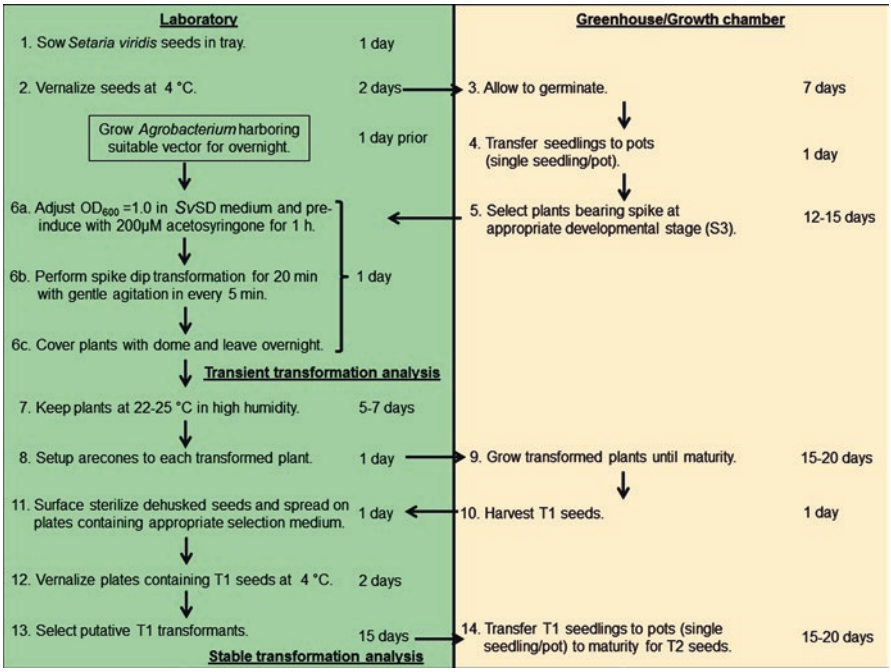


**Fig. 21.2** Flow chart for timeline of *S. viridis* spike-dip transformation procedure. *Sv*SD, *S. viridis* spike dip

resulted in no transformation events. The pH of the medium was adjusted to 5.8 and the surfactant Silwet L-77 (0.025 %) was used. The simple *Sv*I and MS media were not suitable for efficient transformation rates (Saha and Blumwald 2016). Clough and Bent (1998) and Bent (2006) reported that 5 % sucrose was suitable for successful transformation of *Arabidopsis* using floral dip. However, no significant difference in the transformation rates was found when only 5 % Sucrose or MS + Sucrose (pH 5.8) were used in the inoculation medium separately, suggesting that pH adjustments may not be a critical factor for successful spike-dip transformation of *S. viridis* (Saha and Blumwald 2016). Supporting this observation, Clough and Bent (1998) also demonstrated that pH adjustments were unnecessary for successful transformation of *Arabidopsis* using floral dip. The highest efficiency of stable transformations was attained at 20 min of dipping in the presence of 0.025 % Silwet L-77 (Saha and Blumwald 2016) (Fig. 21.2).

## 21.5 Applicable Binary Vector and *Agrobacterium* Strain for Spike-Dip Transformation

Five reporter genes, three [*β-glucuronidase (GUS)* and *β-glucuronidase plus* (*GUSplus*)] for histochemical GUS assay and two [*Aequoria victoria* green fluorescent protein (*GFP*) and *Discosoma* sp. red fluorescent protein (*DsRed*)] for

fluorescent expression assay, were used for transient as well as stable transformations (Fig. 21.1b). The *GUS* reporter gene from *Escherichia coli* of pCAMBIA1201 vector includes an intron from the castor bean catalase gene within the coding sequence driven by *cauliflower mosaic virus 35S* (*CaMV35S*) promoter, ensuring that the constitutive expression of GUS activity was derived from plant cells and not from expression by residual *A. tumefaciens* cells. The pH7m24GWp35S-GUSplus and pH7m24GWpUbi-GUSplus vectors containing the *GUSPlus* gene were originally isolated from a *Staphylococcus* species under the control of either the *CaMV35S* or an intron-interrupted maize *ubiquitin* (*Ubi*) promoter, respectively. Chen et al. (2010) demonstrated that the *GUSplus* reporter gene was useful for high-throughput transient gene expression in *Panicum virgatum* and is more stable at higher temperatures and in fixatives than the *GUS* gene. The *GFP* and the *DsRed* reporter genes were under the control of *CaMV35S* and *Ubi* promoters in pH7m24GW35Sp-GFP and pGWB17-UbiDsRED-UbiBAR binary vectors, respectively. These vectors were reported to be suitable for *Agrobacterium* transformations (Jach et al. 2001; Sheen et al. 1995). The *Ubi* promoter was found to be useful for expression of transgenes in the spike-dip transformation of *S. viridis* (Saha and Blumwald 2016). The maize *Ubi* promoter has been extensively employed for the enhanced constitutive expression of target genes in monocot cereals (Toki et al. 1992; Hiei et al. 2014). The choice and the selection of antibiotic concentration are also important considerations when using this protocol. *Hygromycin phosphotransferaseII* (*HptII*) and bialaphos *(Bar)* are plant selection marker genes that confer resistance to hygromycin B and phosphinothricin (PPT) (the active ingredient in the broad-spectrum herbicide basta), respectively. These selection genes provide relatively low escapes and were found to be suitable for transformed plant selection in this protocol. Hygromycin B (30 mg/L) and basta (3 mg/L) concentrations were critical and should not be as high as to be lethal to low copy number T-DNA transformants (Wilmink and Dons 1993).

*Agrobacterium* strains differ in their virulence to infect target plant species. Out of four *Agrobacterium tumefaciens* strains used, EHA105 (Hood et al. 1993) and AGL1 (Lazo et al. 1991) showed a superior transformation rate, while GV3101 (Koncz and Schell 1986) and LBA4404 (Hoekema et al. 1983) were unable to transform plants at high efficiency (Saha and Blumwald 2016). Both AGL1 (Lazo et al. 1991) and EHA105 (Hood et al. 1993) contain succinamopine type Ti plasmid with C58 origin (Hamilton and Fall 1971) and were shown previously to be suitable for monocot transformation (Chen et al. 2010; Wu et al. 2008; Saha et al. 2006b; Zale et al. 2009; Eck and Swartwood 2015). The effect of the *Agrobacterium* growth phase on stable transformation using the strain EHA105 demonstrated low transformation efficiency at the lag phase ($OD_{600}=0.2$). However, higher transformation rates were seen at the late log phase ($OD_{600}=1.0$) and significant lower rates were obtained at the stationary phase ($OD_{600}=1.5$). Although, *Agrobacterium* at early or late log phases were equally effective for transient transformation, the late log phase was more efficient than the early log phase for stable transformation of *S. viridis* (Saha and Blumwald 2016). *Agrobacterium* at the stationary phase induced severe spike yellowing, wilting, and bleaching.

## 21.6 *Agrobacterium*-Mediated Spike-Dip Transformation Method of *S. viridis*

A single colony of *A. tumefaciens* strain harboring the binary vector was cultured overnight in 5 mL liquid YEB medium (Vervliet et al. 1975) one day before dipping, containing appropriate antibiotics, at 28 °C with an agitation of 250 rpm. The following day, 500 μL of this starter culture was added to 50 mL of culture medium containing 5 g/L tryptone, 2.5 g/L yeast extract, 5 g/L NaCl, 5 g/L mannitol, 100 mg/L, $MgSO_4$, 250 mg/L $K_2HPO_4$, 1.2 g/L glutamic acid, 15 g/L sucrose, pH 7.2 and antibiotics, and the bacteria was grown for 16 h at the same conditions as mentioned earlier. The next day, *Agrobacterium* culture density was adjusted to $OD_{600} = 1.0$ by resuspending in 40 mL *Sv*SD medium supplemented with 200 μM acetosyringone (PhytoTechnology Laboratories, Overland Park, KS) in a 50 mL falcon tube and preinduced by shaking at 180 rpm for 1 h. Prior to spike-dip, plants with single primary tiller bearing a spike at S3 were identified and the remaining secondary spikes were clipped off (Fig. 21.1a). The tillers were labeled for appropriate constructs to transform and the spikes were gently unsheathed (Fig. 21.1a, b). The spikes were immersed into the *Agrobacterium* culture for 20 min with occasional gentle agitation at 5 min intervals (Fig. 21.1c), before placing them under a plastic dome to retain humidity for 24 h, either in a growth chamber or in a culture room under the low light intensity. Next day, plants were returned at 25 °C for 16 h day/8 h night photoperiod with 50 % relative humidity in the growth chamber or in the growth room where they grew for 5–7 days. Plants were placed inside the arecones (arabase and aratubes, Lehle Seeds) (Fig. 21.1d) and transient transformation assays were performed during this time on 3 or 5 days postdipped (DPD) spikes (Fig. 21.1e, f). Finally, plants were grown under greenhouse conditions as mentioned before for 2–4 weeks until maturity. When spikes turned brown and dry, and set seeds, seeds were harvested by gently removing the arecones and collected in microfuge tubes and stored at 4 °C under desiccation.

Mature seeds from T1 and subsequent generations were dehusked, and surface sterilized using 10 % of commercial bleach with 0.1 % Tween 20 for 5 min followed by 3–5 washings in sterile deionized water and blotted dry on sterile filter papers for 5 min. Expression of reporter genes in mature T1 seeds (Fig. 21.1g, h) and seedling (Fig. 21.1i) was monitored. Seeds (20–25) were transferred to plates containing half strength MS medium fortified with either 30 mg/L hygromycin (PhytoTechnology Laboratories, Shawnee Mission, KS) or 3 mg/L bialaphos (PhytoTechnology Laboratories) as described earlier (Saha et al. 2006b). Petri plates and lids were sealed with Parafilm tape and kept at 4 °C for 2 days for seed vernalization. Plates were then kept horizontally and incubated in the growth chamber (Percival Scientific, Perry, IA) at 25 °C with 16 h day/8 h night photoperiod and 50 % relative humidity. Putative transformants were identified as hygromycin-resistant seedlings with well-developed green leaves showing profuse rooting after 10 days postgermination (DPG) on hygromycin selection (Fig. 21.1j). Selected putative transformants were transferred to soil in pots and grown under greenhouse conditions until maturity. The percentage (%) of transformation efficiencies were calculated as (number of hygromycin-resistant seedlings)/(total number

seeds tested)×100. An efficiency of 0.5–0.8% was obtained using *Agrobacterium* strain EHA105 using all three water-deficit and high-temperature tolerant genotypes (Table 21.1) based on the selection of T1 seeds on hygromycin plates (Fig. 21.1j). The germination of false positives during the selection of T1 seeds on hygromycin-containing plates was observed after 7 DPG (Saha and Blumwald 2016). These seedlings grew slowly and turned yellow to brown and eventually died. On the other hand, the hygromycin-tolerant transformed seedlings remained green with well-developed roots.

## 21.7 Molecular and Segregation Analyses of Transgenic Plants

Primers for PCR and quantitative real-time PCR (qRT-PCR) were designed as described before (Kim et al. 2015) and given in Table 21.2. Genomic DNA isolation from transformed and untransformed control plant tissue was carried out following a CTAB extraction method as reported earlier (Ray et al. 2006). The stable integration of transgenes in the *S. viridis* genome was carried out using PCR with gene specific primers (Table 21.2) following the amplification program reported by Saha et al. (2007). Agarose gel electrophoresis analysis of the PCR products obtained from randomly selected independent transgenic plants revealed stable integration of T-DNA gene cassettes into the genome of *S. viridis* (Saha and Blumwald 2016).

The copy number of the integrated T-DNA was confirmed by Southern blot hybridization using 10 μg genomic DNA from transformed and untransformed control plants. Restriction digestion of genomic DNA, gel electrophoresis, transfer, probe preparation, hybridization, and autoradiography were essentially carried out according to procedures described earlier (Saha et al. 2006a). Analysis of autoradiograms showed that the majority of transgenic lines had single copy T-DNA integration (Saha and Blumwald 2016).

Segregation analyses carried out according to Dutta et al. (2005) showed a clear monogenic 3:1 ratio of resistant: susceptible of T2 progeny plants derived from self-fertilized single copy T-DNA integrated T1 plants (Saha and Blumwald 2016).

Total RNA isolation and the first strand cDNA synthesis were carried out according to Saha and Blumwald (2014). The qRT-PCR reactions for determining the fold change ($2^{-\Delta\Delta CT}$) of transgenes expression were conducted following Saha et al. (2013). Expression analyses showed significant transcript levels of inherited transgenes in the selected T2 generation (Saha and Blumwald 2016).

The development of this *in planta Agrobacterium*-mediated spike-dip transformation will facilitate functional genomics studies in an emerging monocot model *S. viridis*.

**Acknowledgements** Authors are thankful to Dr. Ellen Tumimbang, Dr. Hiromi Tajima, Elham Abed, and Kevin Abernathy for technical support. This work was funded by the United States Agency for International Development (USAID) to support the Feed the Future Innovation Lab for Climate-Resilient Millet under the Grant No. APS M/OAA/GRO/EGAS-11-002011.

**Table 21.2** Primers used for molecular analysis of *Setaria viridis* transgenic plants generated through spike-dip transformation method

| Gene | Description | Primer | Sequence (5′–3′) | Size | Use |
|---|---|---|---|---|---|
| *35S::GUS* | β-glucuronidase (GUS) coding gene | *35S*_pF | GGAAGGTGGCTCCTACAAATG | 666 | PCR |
| | | *GUS*_pR | GAATATCTGCATCGGCGAACTG | | |
| | | *GUS*_qF | GAATACGGCGTGGATACGTTAG | 105 | qRT-PCR |
| | | *GUS*_qR | GATCAAAGACGCGGTGATACA | | |
| *GUSplus* | β-glucuronidase plus (GUSplus) coding gene | *GUS*+_pF | CTGGAAGAGAAGTGGTACGAAAG | 545 | PCR |
| | | *GUS*+_pR | GCCTTGAAAGTCCACCGTATAG | | |
| | | *GUSplus*_qF | CGAGCAATGTGATGGATTTC | 100 | qRT-PCR |
| | | *GUSplus*_qR | ATCCGCAAGACGCATCAA | | |
| *GFP* | Green fluorescent protein (GFP) coding gene | *GFP*_pF | ATGGTAGATCTGACTAGTAAAGGAG | 752 | PCR |
| | | *GFP*_pR | CACGTGGTGGTGGTGGTGGTGG | | |
| | | *GFP*_qF | TCCACACAATCTGCCCTTTC | 123 | qRT-PCR |
| | | *GFP*_qR | GGTGGTGGCTAGCTTTGTATAG | | |
| *DsRed* | DsRed fluorescent protein coding gene | *DsRed*_pF | ATGGCGCGCTCCTCCAAGAACG | 680 | PCR |
| | | *DsRed*_pR | CTACAGGAACAGGTGGTGGCGG | | |
| | | *DsRed*_qF | CACTACCTGGTGGAGTTCAAG | 116 | qRT-PCR |
| | | *DsRed*_qR | GATGGTGTAGTCCTCGTTGTG | | |
| *HptII* | Hygromycin phosphotransferase coding gene | *HptII*_pF | ACTCACCGCGACGTCTGTCG | 578 | PCR |
| | | *HptII*_pR | TGCGGCCATTGTCCGTCAGG | | |
| | | *HptII*_qF | CGATGCAAAGTGCCGATAAAC | 101 | qRT-PCR |
| | | *HptI*I_qR | GCTTTCAGCTTCGATGTAGGA | | |
| *Bar* | Bar coding gene | *Bar*_pF | ATGAGCCCAGAACGACGCCCGG | 500 | PCR |
| | | *Bar*_pR | TCAGATTTCGGTGACGGGCAGG | | |
| | | *Bar*_qF | GTCGTCCGTCCACTCCT | 81 | qRT-PCR |
| | | *Bar*_qR | CACCATCGTCAACCACTACATC | | |
| *β-TUB* | β-Tubulin gene | *βTUB*_pF | TGCTGCTGCTGTAATCTGGATC | 759 | PCR |
| | | *βTUB*_pR | TCCACGAAGTAGGAGGAGTTC | | |
| *ACT* | Actin gene | *ACT*_qF | CCTCTTCCAGCCATCTTTCAT | 118 | qRT-PCR |
| | | *ACT*_qR | GAGGACGATGTTGCCGTATAG | | |

*F* forward, *R* reverse, *p* PCR, *q* qRT-PCR, *bp* base pair

## References


Bastaki NK, Cullis CA. Floral-dip transformation of flax (Linum usitatissimum) to generate transgenic progenies with a high transformation rate. J Vis Exp. 2014;94:1–10. doi:10.3791/52189.

Bechtold N, Pelletier G. In planta Agrobacterium-mediated transformation of adult Arabidopsis thaliana plants by vacuum infiltration. Methods Mol Biol. 1998;82:259–66.

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye C-Y, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM. Reference genome sequence of the model plant Setaria. Nat Biotechnol. 2012;30(6):555–61. doi:10.1038/nbt.2196.

Bent A. Arabidopsis thaliana floral dip transformation method. Methods Mol Biol. 2006;343:87–103. doi:10.1385/1-59745-130-4:87.

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu XG, Kellogg E, Van Eck J. Setaria viridis: a model for C4 photosynthesis. Plant Cell. 2010;22(8):2537–44. doi:10.1105/tpc.110.075309.

Chen X, Equi R, Baxter H, Berk K, Han J, Agarwal S, Zale J. A high-throughput transient gene expression system for switchgrass (Panicum virgatum L.) seedlings. Biotechnol Biofuels. 2010;3:1–10. doi:10.1186/1754-6834-3-9.

Clough SJ, Bent AF. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 1998;16(6):735–43. doi:10.1046/j.1365-313x.1998.00343.x.

Curtis I, Nam H. Transgenic radish (Raphanus sativus L. longipinnatus Bailey) by floral-dip method—plant development and surfactant are important in optimizing transformation efficiency. Transgenic Res. 2001;10(4):363–71. doi:10.1023/A:1016600517293.

Desfeux C, Clough SJ, Bent AF. Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method. Plant Physiol. 2000;123(3):895–904.

Dutta I, Saha P, Majumder P, Sarkar A, Chakraborti D, Banerjee S, Das S. The efficacy of a novel insecticidal protein, Allium sativum leaf lectin (ASAL), against homopteran insects monitored in transgenic tobacco. Plant Biotechnol J. 2005;3(6):601–11. doi:10.1111/j.1467-7652.2005.00151.x.

Eck J, Swartwood K. Setaria viridis. In: Wang K, editor. Agrobacterium protocols, vol. 1. New York: Springer; 2015. p. 57–67. doi:10.1007/978-1-4939-1695-5_5.

Hamilton RH, Fall MZ. The loss of tumor-initiating ability in Agrobacterium tumefaciens by incubation at high temperature. Experientia. 1971;27(2):229–30.

Hiei Y, Ishida Y, Komari T. Progress of cereal transformation technology mediated by Agrobacterium tumefaciens. Front Plant Sci. 2014;5:628. doi:10.3389/fpls.2014.00628.

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA. A binary plant vector strategy based on separation of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid. Nature. 1983;303(5913):179–80.

Hood E, Gelvin S, Melchers L, Hoekema A. New Agrobacterium helper plasmids for gene transfer to plants. Transgenic Res. 1993;2(4):208–18. doi:10.1007/BF01977351.

Jach G, Binot E, Frings S, Luxa K, Schell J. Use of red fluorescent protein from Discosoma sp. (dsRED) as a reporter for plant gene expression. Plant J. 2001;28(4):483–91. doi:10.1046/j.1365-313X.2001.01153.x.

Kim H-Y, Saha P, Farcuh M, Li B, Sadka A, Blumwald E. RNA-seq analysis of spatiotemporal gene expression patterns during fruit development revealed reference genes for transcript normalization in plums. Plant Mol Biol Rep. 2015;33(6):1634–49. doi:10.1007/s11105-015-0860-3.

Koncz C, Schell J. The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Mol Gen Genet. 1986;204(3):383–96. doi:10.1007/BF00331014.

Lazo GR, Stein PA, Ludwig RA. A DNA transformation-competent Arabidopsis genomic library in Agrobacterium. Biotechnology. 1991;9(10):963–7.

Lee MW, Yang Y. Transient expression assay by agroinfiltration of leaves. Methods Mol Biol. 2006;323:225–9. doi:10.1385/1-59745-003-0:225.

Liu X, Brost J, Hutcheon C, Guilfoil R, Wilson A, Leung S, Shewmaker C, Rooke S, Nguyen T, Kiser J, De Rocher J. Transformation of the oilseed crop Camelina sativa by Agrobacterium-mediated floral dip and simple large-scale screening of transformants. In Vitro Cell Dev Biol Plant. 2012;48(5):462–8. doi:10.1007/s11627-012-9459-7.

Martins PK, Nakayama TJ, Ribeiro AP, Cunha BADB, Nepomuceno AL, Harmon FG, Kobayashi AK, Molinari HBC. Setaria viridis floral-dip: a simple and rapid Agrobacterium-mediated transformation method. Biotechnol Rep. 2015a;6:61–3. doi:10.1016/j.btre.2015.02.006.

Martins PK, Ribeiro AP, Cunha BADB, Kobayashi AK, Molinari HBC. A simple and highly efficient Agrobacterium-mediated transformation protocol for Setaria viridis. Biotechnol Rep. 2015b;6:41–4. doi:10.1016/j.btre.2015.02.002.

Mohanty A, Sarma NP, Tyagi AK. Agrobacterium-mediated high frequency transformation of an elite indica rice variety Pusa Basmati 1 and transmission of the transgenes to R2 progeny. Plant Sci. 1999;147(2):127–37. doi:10.1016/S0168-9452(99)00103-X.

Mu G, Chang N, Xiang K, Sheng Y, Zhang Z, Pan G. Genetic transformation of maize female inflorescence following floral dip method mediated by Agrobacterium. Biotechnology. 2012; 11(3):178–83.

Murashige T, Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant. 1962;15(3):473–97. doi:10.1111/j.1399-3054.1962.tb08052.x.

Ray T, Dutta I, Saha P, Das S, Roy SC. Genetic stability of three economically important micropropagated banana (Musa spp.) cultivars of lower Indo-Gangetic plains, as assessed by RAPD and ISSR markers. Plant Cell Tiss Organ Cult. 2006;85(1):11–21. doi:10.1007/s11240-005-9044-4.

Sage RF, Zhu X-G. Exploiting the engine of C4 photosynthesis. J Exp Bot. 2011;62(9):2989–3000. doi:10.1093/jxb/err179.

Saha P, Blumwald E. Assessing reference genes for accurate transcript normalization using quantitative real-time PCR in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. PLoS One. 2014;9(8):e106308. doi:10.1371/journal.pone.0106308.

Saha P, Blumwald E. Spike-dip transformation of Setaria viridis. Plant J. 2016;86(1):89–101. doi:10.1111/tpj.13148.

Saha P, Dasgupta I, Das S. A novel approach for developing resistance in rice against phloem limited viruses by antagonizing the phloem feeding hemipteran vectors. Plant Mol Biol. 2006a;62(4–5):735–52. doi:10.1007/s11103-006-9054-6.

Saha P, Majumder P, Dutta I, Ray T, Roy SC, Das S. Transgenic rice expressing Allium sativum leaf lectin with enhanced resistance against sap-sucking insect pests. Planta. 2006b;223(6):1329–43. doi:10.1007/s00425-005-0182-z.

Saha P, Chakraborti D, Sarkar A, Dutta I, Basu D, Das S. Characterization of vascular-specific RSs1 and rolC promoters for their utilization in engineering plants to develop resistance against hemipteran insect pests. Planta. 2007;226(2):429–42. doi:10.1007/s00425-007-0493-3.

Saha P, Ray T, Tang Y, Dutta I, Evangelous NR, Kieliszewski MJ, Chen Y, Cannon MC. Self-rescue of an EXTENSIN mutant reveals alternative gene expression programs and candidate proteins for new cell wall assembly in Arabidopsis. Plant J. 2013;75(1):104–16. doi:10.1111/tpj.12204.

Saha P, Sade N, Arzani A, Wilhelmi MMR, Coe KM, Li B, Blumwald E. Effects of abiotic stress on physiological plasticity and water use on Setaria viridis. Plant Sci. 2016;251:128–138. doi:org/10.1016/j.plantsci.2016.06.011

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K,

Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh C-T, Emrich SJ, Jia Y, Kalyanaraman A, Hsia A-P, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia J-M, Deragon J-M, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK. The B73 maize genome: complexity, diversity, and dynamics. Science. 2009;326(5956):1112–5.

Sheen J, Hwang S, Niwa Y, Kobayashi H, Galbraith DW. Green-fluorescent protein as a new vital marker in plant cells. Plant J. 1995;8(5):777–84. doi:10.1046/j.1365-313X.1995.08050777.x.

Toki S, Takamatsu S, Nojiri C, Ooba S, Anzai H, Iwata M, Christensen AH, Quail PH, Uchimiya H. Expression of a maize ubiquitin gene promoter-bar chimeric gene in transgenic rice plants. Plant Physiol. 1992;100(3):1503–7.

Trieu AT, Burleigh SH, Kardailsky IV, Maldonado-Mendoza IE, Versaw WK, Blaylock LA, Shin H, Chiou T-J, Katagi H, Dewbre GR, Weigel D, Harrison MJ. Transformation of Medicago truncatula via infiltration of seedlings or flowering plants with Agrobacterium. Plant J. 2000;22(6):531–41. doi:10.1046/j.1365-313x.2000.00757.x.

Vervliet G, Holsters M, Teuchy H, Van Montagu M, Schell J. Characterization of different plaque-forming and defective temperate phages in Agrobacterium. J Gen Virol. 1975;26(1):33–48. doi:10.1099/0022-1317-26-1-33.

Wang QM, Wang L. An evolutionary view of plant tissue culture: somaclonal variation and selection. Plant Cell Rep. 2012;31(9):1535–47. doi:10.1007/s00299-012-1281-5.

Wilmink A, Dons JJM. Selective agents and marker genes for use in transformation of monocotyledonous plants. Plant Mol Biol Rep. 1993;11(2):165–85. doi:10.1007/BF02670474.

Wu H, Doherty A, Jones HD. Efficient and rapid Agrobacterium-mediated genetic transformation of durum wheat (Triticum turgidum L. var. durum) using additional virulence genes. Transgenic Res. 2008;17(3):425–36. doi:10.1007/s11248-007-9116-9.

Yasmeen A, Mirza B, Inayatullah S, Safdar N, Jamil M, Ali S, Choudhry MF. In planta transformation of tomato. Plant Mol Biol Rep. 2009;27(1):20–8. doi:10.1007/s11105-008-0044-5.

Ye G-N, Stone D, Pang S-Z, Creely W, Gonzalez K, Hinchee M. Arabidopsis ovule is the target for Agrobacterium in planta vacuum infiltration transformation. Plant J. 1999;19(3):249–57. doi:10.1046/j.1365-313X.1999.00520.x.

Zale JM, Agarwal S, Loar S, Steber CM. Evidence for stable transformation of wheat by floral dip in Agrobacterium tumefaciens. Plant Cell Rep. 2009;28(6):903–13. doi:10.1007/s00299-009-0696-0.

# Index





A. Doust, X. Diao (eds.), *Genetics and Genomics of Setaria*, Plant Genetics and Genomics: Crops and Models 19, DOI 10.1007/978-3-319-45105-3

**G**