REVIEW

Initiation of Setaria as a model plant

Xianmin DIAO (✉)[1], James SCHNABLE[1], Jeffrey L. BENNETZEN[2], Jiayang LI (✉)[3]

1 Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
2 Departments of Genetics, University of Georgia, Athens, GA30602, USA and Germplasm Bank of Wild Species in Southwestern China,
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650000, China
3 Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract Model organisms such as Arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*) have proven essential for efficient scientific discovery and development of new methods. With the diversity of plant lineages, some important processes such as C_4 photosynthesis are not found in either Arabidopsis or rice, so new model species are needed. Due to their small diploid genomes, short life cycles, self-pollination, small adult statures and prolific seed production, domesticated foxtail millet (*Setaria italica*) and its wild ancestor, green foxtail (*S. viridis*), have recently been proposed as novel model species for functional genomics of the Panicoideae, especially for study of C_4 photosynthesis. This review outlines the development of these species as model organisms, and discusses current challenges and future potential of a Setaria model.

Keywords Setaria, foxtail millet, C_4 photosynthesis, model organism

1 Introduction: the leading role of model organisms

Model organisms are significant in enabling new scientific discovery and the development of new technology. Discoveries in early genetic models such as, pea (*Pisum sativum*), fruit fly (*Drosophila melanogaster*) and maize (*Zea mays*), laid the foundation for modern genetics. Research with these species has been responsible for the discovery of everything from the basic laws of genetics and heritability to the exceptions to those laws, such as transposable elements, and to the pioneering of the wholly new field of cytogenetics. This early understanding of genetics and inheritance led to the establishment of dramatically improved breeding methods using hybridization and single line selection. In the genomics era, new models such as Arabidopsis and rice were used to discover genetic regulatory networks, such as the ABC model of floral development. The technological progress of the genomics era, first elaborated in model species, has enabled novel approaches to breeding including marker-assisted selection and breeding by molecular design.

Arabidopsis has been developed extensively as a model for plant molecular genetics and functional genomics, and this organism was especially suitable for answering questions about fundamental mechanisms that are conserved across all plant species. However, flowering plants are an extremely diverse group of species, with many traits and systems of particular importance to humans that are found only in specific plant lineages, such as C_4 photosynthesis. Characterizing these lineage-specific systems requires the identification and development of new model species.

Cereal crops, members of the grass family, contribute close to 70% of the calories consumed by humans around the world, and are only distantly related to the eudicot clade, which contains Arabidopsis. Many grass species have also evolved to use the C_4 photosynthetic cycle, which is more energetically productive and uses both nitrogen and water more efficiently than the C_3 photosynthetic cycle found in most green plants [1]. Within the grasses, rice and Brachypodium (*Brachypodium distachyon*) are also being developed as model systems for the grasses. However, rice and Brachypodium are both C_3 plants, so they are not suitable for many C_4-related functional analyses. In order to both characterize the evolutionary innovations responsible for C_4 photosynthesis and provide genomic insight into crop improvement for the many important crop species found within the panicoid grasses, a new genomic and genetic model species is required.

2   Initiation of the Setaria model

The grass genus *Setaria*, part of the tribe Paniceae in the subfamily Panicoideae of the family Poaceae, contains approximately 125 species world-wide. Species in this genus are found in tropical, sub-tropical and temperate regions, and include both crop and weed species with different life cycles and ploidy levels [2]. So far, seven genome types have been identified in the genus and their genetic and structural relationships have been investigated by genomic *in situ* hybridization (GISH) and molecular marker analyses [3]. Two species of *Setaria*, domesticated foxtail millet (*S. italica*) and its wild ancestor green foxtail (*S. viridis*), exhibit numerous properties (for instance, C_4 photosynthesis) that make them ideal models for functional genomics studies in the Panicoid grasses. Both species are diploid and have tractably small genomes (490 megabases, about one fifth the size of the maize genome), short generation times (50–90 days), small stature and prolific seed production. These two *Setaria* species show excellent genetic colinearity, because foxtail millet is believed to have been domesticated just within the last few thousand years from the weedy green foxtail [4]. Given their hybrid fecundity, foxtail millet and green foxtail could be viewed as a single species with two botanical names. With a high quality reference genome sequence [4] and a high-density haplotype map of genome variation [5] and other genomic data, the accumulated genomic information for foxtail millet and green foxtail has reached the level where this species pair can now truly be considered a novel model system. Several publications, foxtail millet: a sequence-driven grass model system by Doust et al. [6], *Setaria viridis*: a model for C_4 photosynthesis by Brutnell et al. [7], and two review papers of using foxtail millet and green foxtail as model [8,9] have helped foster enthusiasm and forward momentum that is leading to the development of *Setaria* as a model system. A note on terminology: although *Setaria* is the name of a genus, Setaria is used here to refer to either the domesticated species, foxtail millet, or the wild ancestor, green foxtail, following the common convention (i.e. *A. thaliana* is called Arabidopsis as a model system and *B. distachyon* is called Brachy or Brachypodium).

To help initiate Setaria as a model, the first International Setaria Genetics Conference (ISGC 2014) was held in Beijing on 10-12 March 2014. The conference brought together 230 scientists and researchers from 9 countries working on different aspects of Setaria, including germplasm collection, phylogenetic analysis, genome variation, mutant library construction, EST and other tool establishment, the biology of C_4 photosynthesis, genetic transformation, using Setaria as model to study sugarcane, foxtail millet breeding, and other aspects of genetics and genomics. An International Setaria Community was constructed during ISGC 2014 and this marked both the birth of a novel model organism and created a roadmap for the future of Setaria.

3   Setaria as model for the Panicoideae crops and C_4 photosynthesis

The species group of the Panicoideae is the biggest sub-family in the grasses, and includes many C_4 species and a large number of important crop species: maize (*Zea mays*), sorghum (*Sorghum bicolor*), sugarcane (*Saccharum officinarum*), pearl millet (*Pennisetum glaucum*), foxtail millet (*Setaria italica*), common millet (*Panicum miliaceum*), switchgrass (*Panicum virgatum*), napiergrass (*Pennisetum purpureum*) and some candidate biofuel crops. Belonging to this subfamily of the grasses with close evolution relationships, a Setaria model will certainly facilitate functional genomics for these crops. Maize is a widely grown species of great economic importance, and a history of genetic investigation going back more than a century. However, in the genomics era, the large genome size of maize combined with its paleopolyploid evolutionary history and the shear size of adult maize plants present challenges for functional genomics research. Sorghum is also considered as a model for C_4 functional genomics study [10], but its adult plant size and longer life cycle made it difficult for lab management. Switchgrass and napiergrass are considered putative biofuel crops, but their polyploid backgrounds make their functional genomics study difficult, so is the situation to common millet, an ancient cereal which is still widely grown in dry land area in the world. The establishment of Setaria as a model will assist in the genetic study of all these species, and many more Panicoid grasses.

In addition to many individual research groups studying Setaria, the C_4 rice consortium, which is funded by the Gates Foundation, and several private biotech companies are currently using Setaria in attempts to decipher the genetic mechanisms underlying C_4 photosynthesis [11,12]. C_4 species are characterized by the Kranz Structure of bundle sheath cells and C_4 acid shuttling-related enzymes such as malate dehydrogenase, phosphoenolpyruvate carboxylase and pyruvate orthophosphate dikinase. C_3 species, such as rice and Brachypodium, share the primary carbon shuttle enzymes with C_4 species of Setaria, sorghum and maize, albeit the amino acid identity of those enzymes between C_4 and C_3 might be low, and it is deduced those enzymes were recently recruited into the C_4 pathway from ancestral C_3 isoforms [4]. Understanding the genes necessary for Kranz structure development in C_4 species will be essential for the deciphering of C_4 genetic and molecular mechanisms, and identification of Kranz structural mutations in C_4 plants is one appropriate approach. The small stature and short life cycle of Setaria make it suitable for high throughput operations for the screening of related mutations.

4   Setaria as a model for the study of plant abiotic stress response

Studies of plant responses to abiotic stresses have long been of great interest to scientist because of growing water shortages and decreasing soil fertility. Foxtail millet seeds require only 26% of their seed weight in water for successful germination, while other cereals require at least that of 45% of their seed weight [13]. The water use efficiency (WUE) of foxtail millet is also higher than other grain crops. Maize requires 470 g of water to produce 1 g of dry biomass and wheat requires 510 g of water, while foxtail millet requires a modest 257 g [13]. Several morphological and physiological adaptations have been reported to be associated with drought tolerance and higher WUE in foxtail millet, including dense and deep root systems, smaller leaf area, and thickening of cell walls [14].

High throughput assessment and screening of drought tolerant foxtail millet genotypes from 17313 accessions was carried out in China by measuring the rate of seedling survival under multiple drought stress treatments. More than two hundred lines were classified into the highest category of drought tolerance, including the reference genome cultivar, Yugu 1 [14]. A rapid and high throughput screening for drought tolerance during germination in foxtail millet was developed which used polyethylene glycol (PEG-6000) to generate osmotic stress [15]. Multiple methods of gene expression profiling have been employed to identify genes related to drought tolerance in foxtail millet, and these methods identified hundreds of significant changes in expression that corresponded to genes involved in metabolism, proteolysis, and signaling [16–19]. However, the detailed genetic mechanisms responsible for variation in drought tolerance among foxtail millet lines remain uncharacterized. Comprehensive analyses of combination with drought tolerant quantitative trait loci (QTL) mapping, expression profiling and mutation screening are needed in the future.

5   Promoting Setaria as a model

Realizing the ideal nature of the Setaria model system, many leading research groups around the world have begun to work with Setaria. Extensive progress has been made in characterizing different aspects of Setaria genetics. In addition to the two genome sequences of foxtail millet released in 2012 [4,20], a third genome sequence of foxtail millet (landrace Daqingjie) and a genome sequence of green foxtail (Accession N10) have also been released [5]. Re-sequencing of foxtail millet landrace Shilixiang was also carried out to identify useful markers for studies of agronomic characters [21]. These sequence data provided the foundation for the development of polymorphic markers, leading to a high density genetic map of Setaria based on SSRs, which will facilitate genetic and genomic study of this model plant [22,23]. Genetic diversity and population structure studies in Chinese green foxtail [24] and domesticated landraces [25] revealed high level of genetic diversity, an average allele number per locus of 33.5 and 20.9, respectively. A high density haplotype map of 85 million single nucleotide polymorphism (SNP) markers was also constructed, which revealed genomic variation among 916 foxtail millet accessions collected from around the world [5]. All these analyses provide the genomic data for the development of the Setaria model system.

Populations of mutant plants can be vital in functional genomic studies, and will be as useful for Setaria as it has been for Arabidopsis and rice. Using the genome sequenced foxtail millet cultivar Yugu 1 as material, an ethylmethanesulphonte (EMS) mutant library was constructed [26]. Different kinds of mutants were identified from the M_2 lines which include sterile flowers, early or late heading, yellow leaves, dwarf stem, small panicle, curled leaves and many kind of panicle morphology variation. Among all these mutants, flower sterility and panicle shape mutations were most frequently obtained. Among the 7380 lines of the library, 316 lines were identified with evident morphological variation. A mutation population created by fast-neutron treatment of foxtail millet was also established with Yugu1 for functional genomics of Setaria [27]. Although the size of these libraries is small, their establishment provides basic material for functional genomic study of foxtail millet, and will help Setaria develop as a model for the grasses.

6   A good chance for genetic improvement of foxtail millet

Foxtail millet is recognized as one of the most ancient domesticated cereals. As a staple food, it contributed to the development of human civilization and is still grown as a staple food, particularly in China and India. Foxtail millet is more tolerant of drought and uses water and fertilizer more efficiently than many cereal crops, and thus may be a key to dealing with a warmer and drier future. However, the breeding technology used in foxtail millet is far behind that of major cereals like rice and maize, and the current yield potential of foxtail millet cultivars is relatively low (about 5–7 t·hm^{-2}, compared to 10–12 t·hm^{-2} for maize). Germplasm diversity analysis, genetic investigation, and functional genomic studies of foxtail millet and green foxtail will lay the foundations for molecular breeding in foxtail millet. Given its limited improvement so far, the genome-facilitated development of foxtail millet breeding technologies are certain to lead to the development of

foxtail millet cultivars with both much higher yields and more desirable food quality characteristics.

7   Present challenges and future work

Roundtable discussions at the ISGC 2014 indicated several key needs for development of Setaria as a model system. These needs/opportunities are described briefly below.

The first essential need for a successful Setaria model is the establishing of a stable and efficient transformation protocol. The efficient transformation systems available in Arabidopsis and rice greatly contributed to the maturation of these two species as model systems. Although there are reports of transformation of foxtail millet [7], many scientists still find transformation to be challenging. Fortunately, there were four papers presented at ISGC 2014 describing the establishment of efficient protocols for transforming of Setaria, and one report of very high transformation rate of 29% [28]. There were also several examples of transgenic success using Setaria to test the function of genes identified in sugarcane and other crops presented at ISGC 2014. Testing different components for optimizing transformation efficiency, such as screening diverse genotypes, as well as testing different types of explants, culture media and conditions, is currently being undertaken. The exchange of progress and discoveries from different efforts will provide a chance to establish a stable and efficient Setaria transformation system.

A second urgent need for development of Setaria as model is for a well-organized online database devoted to Setaria genetic and genomic information, along the lines of TAIR for Arabidopsis (www.arabidopsis.org) or MaizeGDB for maize (www.maizegdb.org). This website could also be used to provide a searchable interface for a sequence-tagged mutant library constructed using tDNA insertions or alternatively a set of resequenced EMS/N-nitrosomethylurea (NMU) /fast neutron mutant lines that would provide an equivalent reverse genetics resource. Although there are reports describing the initial generation of EMS mutant populations of Setaria [26], and other mutant populations are in development, it will take a concerted effort to index the known mutations in these lines and make these data widely available to the community.

A third critical need for the Setaria community is the establishment of a core set of cultivars and populations that will be studied by individual laboratories. Several reports describe research investigating genetic diversity in Setaria [24,25], while others at ISGC 2014 discussed tool development in a key set of lines (for instance, *S. viridis* Accession A10) [4]. Other groups described the construction of key mapping populations [29]. Once a core set of genetic materials are settled upon by the Setaria community, comparison of genetic results and tool development will be much more efficient and productive.

8   Prospective

In addition to serving as a genetic model for the Panicoid crops and C_4 photosynthesis, Setaria promises to be an excellent model genetic system for understanding the processes of domestication and variety improvement, for study of the evolution of abiotic stress tolerance, for comparative mapping approaches that span multiple species, for investigating the adaptive evolution of the grasses, and for investigating traits essential for biomass production by novel bioenergy crops. The establishment of efficient transformation methods, the construction of a well-organized Setaria website, and, most important, the launching of an international Setaria community will ensure a prominent and successful future for Setaria as a model system that will enable groundbreaking scientific discoveries, from basic questions in plant development to applied crop breeding and crop improvement.

Compliance with ethics guidelines   Xianmin Diao, James Schnable, Jeffrey L. Bennetzen and Jiayang Li declare that they have no conflict of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by the any of the authors.

References

1. Sage R F. The evolution of C_4 photosynthesis. *New Phytologist*, 2004, 161(2): 341–370

2. Rominger J M. Taxonomy of *Setaria* (Gramineae) in North America. In: Illinois Biol Monogr, volume 29. Edited by Horsfall WR, Delevoryas T, De Moss RD, Kruidenier FJ, and Taylor AB. Urbana: University of Illinois Press; 1962, 100–108

3. Zhao M, Zhi H, Doust A N, 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(1): 244

4. Bennetzen J L, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli A C, Estep M, Feng L, Vaughn J N, 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 P C, Panaud O, Kellogg E A, Brutnell T P, Doust A N, Tuskan G A, Rokhsar D, Devos K M. Reference genome sequence of the model plant Setaria. *Nature Biotechnology*, 2012, 30(6): 555–561

5. 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*). *Nature Genetics*, 2013, 45(8): 957–961

6. Doust A N, Kellogg E A, Devos K M, Bennetzen J L. Foxtail millet: a sequence-driven grass model system. *Plant Physiology*, 2009, 149 (1): 137–141

7. Brutnell T P, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X G, Kellogg E, Van Eck J. *Setaria viridis*: a model for C$_4$ photosynthesis. *Plant Cell*, 2010, 22(8): 2537–2544

8. Li P, Brutnell T P. *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. *Journal of Experimental Botany*, 2011, 62(9): 3031–3037

9. Lata C, Gupta S, Prasad M. Foxtail millet: a model crop for genetic and genomic studies in bioenergy grasses. *Critical Reviews in Biotechnology*, 2013, 33(3): 328–343

10. Wang Y H, Upadhyaya H D, Burrell A M, Sahraeian S M, Klein R R, Klein P E. Genetic structure and linkage disequilibrium in a diverse, representative collection of the C$_4$ model plant, *Sorghum bicolor. G3 (Bethesda)*, 2013, 3(5): 783–93

11. von Caemmerer S, Quick W P, Furbank R T. The development of C$_4$ rice: current progress and future challenges. *Science*, 2012, 336 (6089): 1671–1672

12. Karki S, Rizal G, Quick W P. Improvement of photosynthesis in rice (*Oryza sativa* L.) by inserting the C$_4$ pathway. *Rice*, 2013, 6(1): 28

13. Gu S. Relationship between foxtail millet growth and environmental factors, In Gu S. ed., *Foxtail Millet Cultivation and Production in China,* Beijing, Chinese Agricultural Press, 1987, 63–71

14. Li Y. Breeding for foxtail millet drought tolerant cultivars (in Chinese). In Li Y ed., *Foxtail Millet Breeding, Beijing,* Chinese Agricultural Press, 1997, 421–446

15. Zhu X, Song Y, Zhao Z, Shi Y, Liu Y, Li Y, Wang T. Methods for identification of drought tolerance at germination period of foxtail millet by osmotic stress. *Journal of Plant Genetic Resources*, 2008, 9: 62–67

16. 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(1): 121–131

17. Lata C, Sahu P P, Prasad M. Comparative transcriptome analysis of differentially expressed genes in foxtail millet (*Setaria italica* L.) during dehydration stress. *Biochemical and Biophysical Research Communications*, 2010, 393(4): 720–727

18. Yi F, Xie S, Liu Y, Qi X, Yu J. Genome-wide characterization of microRNA in foxtail millet (*Setaria italica*). *BMC Plant Biology*, 2013, 13(1): 212

19. 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 Molecular Biology*, 2013, 83(4-5): 459–473

20. 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. *Nature Biotechnology*, 2012, 30(6): 549–554

21. Bai H, Cao Y, Quan J, Dong L, Li Z, Zhu Y, Zhu L, Dong Z, Li D. Identifying the genome-wide sequence variations and developing new molecular markers for genetics research by re-sequencing a Landrace cultivar of foxtail millet. *PLoS ONE*, 2013, 8(9): e73514

22. 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 and characterization of highly poly-morphic SSR (Simple Sequence Repeat) markers through genome-wide microsatellite variants analysis in Foxtail millet [*Setaria italica* (L.) P. Beauv.]. *BMC Genomics*, 2014, 15: 78

23. Muthamilarasan M, Venkata Suresh B, Pandey G, Kumari K, Parida S K, Prasad M. Development of 5123 intron-length polymorphic markers for large-scale genotyping applications in foxtail millet. *DNA Research*, 2014, 21(1): 41–52

24. 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. *Journal of Experimental Botany*, 2013, 64(12): 3645–3656

25. Wang C F, Jia G Q, Zhi H, Niu Z G, Chai Y, Li W, Wang Y F, Li H Q, Lu P, Zhao B H, Diao X M. Genetic diversity and population structure of Chinese foxtail millet [*Setaria italica* (L.) Beauv.] landraces. *G3 (Bethesda)*, 2012, 2(7): 769–777.

26. Diao X, Li W, Zhi H, Jia G, Ge Y, Chai Y, Li J. Construction of an EMS mutation library for foxtail millet functional genomics. The first international Setaria genetics conference abstracts, Beijing, 2014, 56

27. Bennetzen J, Chaluvadi S, Feng L, Wang H. Setaria genome structure and evolution. The first international Setaria genetics conference abstracts, Beijing, 2014, 6

28. Martins P K, Dias B B A, Ribeiro A P, Kobayashi A K, Molinari H B C. *Setaria viridis*: a tool for functional gene analysis in sugancane. The first international Setaria genetics conference abstracts, Beijing, 2014, 19

29. Zhi H, Jia G, Niu Z, Liu X, Ge Y, Chai Y, Diao X. Construction of an RIL population and segment introgression lines via interspecific cross between *Setaria italica* and *S. viridis*. The first international Setaria genetics conference abstracts, Beijing, 2014, 50