Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

Brachypodium distachyon genomics for sustainable food and fuel production

Michael W Bevan[1], David F Garvin[2] and John P Vogel[3]

Grass crops are the most important sources of human nutrition, and their improvement is centrally important for meeting the challenges of sustainable agriculture, for feeding the world's population and for developing renewable supplies of fuel and industrial products. We describe the complete sequence of the compact genome of *Brachypodium distachyon* (Brachypodium) the first pooid grass to be sequenced. We demonstrate the many favorable characteristics of Brachypodium as an experimental system and show how it can be used to navigate the large and complex genomes of closely related grasses. The functional genomics and other experimental resources that are being developed will provide a key resource for improving food and forage crops, in particular wheat, barley and forage grasses, and for establishing new grass crops for sustainable energy production.

Addresses
[1] John Innes Centre, Colney Lane, Norwich NR4 7UJ, UK
[2] USDA-ARS Plant Science Research Unit and Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA
[3] USDA-ARS Western Regional Research Center, Albany, CA 94710, USA

Corresponding author: Bevan, Michael W (michael.bevan@bbsrc.ac.uk)

Current Opinion in Biotechnology 2010, **21**:211–217

This review comes from a themed issue on
Plant biotechnology
Edited by Antonio Molina and Jim Haseloff

Available online 31st March 2010

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2010.03.006

Introduction

The grass family (Poaceae) is the fourth largest plant family in the world, with over 10 000 species distributed widely across the earth [1]. The top four global agricultural commodities by quantity are grass crops (sugarcane, maize, rice, wheat) [2]. Cow's milk, the sole animal product in the top 10 agricultural commodities by quantity [2], largely comes from animals fed by grasses. Thus, grasses are centrally important for human existence by directly or indirectly serving as the primary source of human nutrition. By 2050 the human population is predicted to increase to 9 billion [3], and an overall *per capita* increase in living standards and therefore consumption of food and energy is also predicted. Furthermore, this increase must be achieved by more sustainable industries. The capture of sunlight and its conversion to chemical energy by photosynthesis is the primary biological process available for sustainable biotechnology, and photosynthesis in land plants remains the largest source of renewable food and energy [4] that can be produced within a sustainable carbon cycle. Primary production from agriculture therefore assumes an important role in the transition to increasingly sustainable food and industrial production methods. Some grass crop species, such as maize, sorghum and sugarcane have evolved C4 photosynthesis, in which CO_2 is concentrated at the sites of carboxylation to increase photosynthetic efficiency. Grass crops are therefore centrally important targets for biotechnological improvement for food and fuel production. In particular the exploitation of a currently untapped resource of grass biomass (primarily lignocellulosic cell walls) is of high interest for sustainable fuel production [5].

Many challenges need to be overcome to create new environmentally sustainable industries based on photosynthesis. Far higher crop yields need to be achieved with fewer inputs, particularly of phosphate and nitrogenous fertilizers, potentially limiting access to water, and the need to conserve biodiversity by confining agricultural production zones, all require hitherto unachieved crop yields. Food and industrial production streams need to be separated and optimized to minimize gearing of food and fuel prices to alleviate economic disadvantage [6]. Global climate change intensifies these production challenges as current crops are poorly adapted to more uncertain and extreme climatic conditions. Consequently significant increases in our knowledge of grass biology that includes systems-level approaches to understanding how biotic and abiotic environments influence yield need to be achieved. Genomics and functional genomics resources are centrally important for this research, and they also directly facilitate biotechnological and genetic improvement through plant breeding. In this review we examine how genomics and functional genomics in the pooid grass *Brachypodium distachyon* can contribute to grass crop improvement.

B. distachyon is a promising model organism for grass research

Arabidopsis is a central model in basic plant biology research for decades owing to its small size, small genome, rapid generation time, and *in planta* transformation potential [7]. Given the broad diversity represented in grass crops, and some fundamental differences in growth and

Table 1

Comparison of model and crop plants

	Brachypodium	Arabidopsis	Rice	Sorghum	Wheat	Maize	Switchgrass
Height (cm)	15–20	15–20	100	170–320	50	155–215	200
Generation time (weeks)	8–12	8–12	30	17	12	14–20	26
Density (plants/m^2)	1000	2000	36	13	50	6	6
Growth requirements	Simple	Simple	Demanding	Simple	Simple	Simple	Simple
Reproduction	Selfing	Selfing	Selfing	Outcrossing: self-compatible	Selfing	Outcrossing: self-compatible	Outcrossing: self-incompatible
Genome size (Mbp)	272[a]	119[a]	382[a]	758[a]	17 000	2048[a]	2400–3200
Ploidy	2X	2X	2X	2X	6X	2X	4X–8X
Cell wall type[b]	Type II	Type I	Type II	Type II	Type II	Type II	Type II

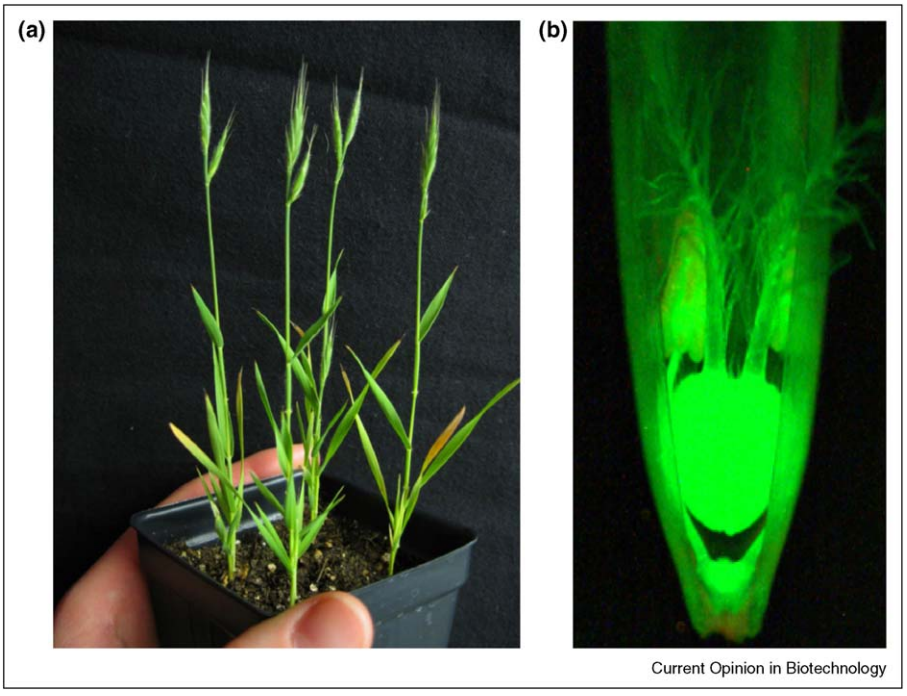
[a] Assembled genome size.
[b] Grass cell walls differ from dicot cell walls in the type of hemicellulose and the amounts of pectin, protein and phenolic compounds [32].

development of grasses compared to dicots such as Arabidopsis, a model grass species that permits investigations comparable to those possible in Arabidopsis would find wide applicability for future grass crop improvement (Table 1).

The small annual (Figure 1) plant species *B. distachyon* belongs to the tribe Brachypodieae, which consists solely of the genus *Brachypodium*. This tribe is located at an intriguing spot in grass phylogeny – basal to the four grass tribes that collectively encompass the vast majority of domesticated cool season cereal grain, forage, and turf crops [8]. The discovery that grass genomes show significant levels of collinearity in gene organization [9] and the appreciation of the full power of focusing research on model systems converged and led to proposals that *B. distachyon* (hereafter Brachypodium), the only annual species in the genus, might have potential as a model grass owing to its small size and small genome (Table 1 and Ref. [10]).

These desirable features of Brachypodium were further described by Draper *et al.* [11], and the potential for developing a 'grass Arabidopsis' drew the attention of

Figure 1



Images of *Brachypodium distachyon* plants. **(a)** Brachypodium plants (genotype Bd21) grown under 23 h day length, 25 days after sowing. Long day lengths permit dense (1 plant per 1.5 cm^2) growth and seed recovery from plants. **(b)** Expression of Green Fluorescent Protein (GFP) in floral organs.

many scientists. Key resources such as community standard genetic stocks, *Agrobacterium tumefaciens*-mediated transformation methods, EST collections, BAC libraries, genetic linkage maps, and segregating populations have since been developed for Brachypodium [12]. The broad interest in Brachypodium and its exceptional promise provided an impetus to sequence the genome as part of the DOE Joint Genome Institute Community Sequencing Program. A detailed analysis of the compact 272 Mb genome has recently been published [13••].

Comparative genomics of the grasses

Three main subfamilies of grasses, the Panicoideae (sorghum, maize), Ehrhartoideae (rice) and Pooideae (wheat, barley) provide the bulk of human and domestic animal nutrition. Although grass genomes vary greatly in size due to expansion of retroelement repeats [14], there is an underlying conserved gene order [9] reflecting their common ancestry and rapid diversification [15]. To date the complete genome sequences of four grass species representing the three most economically important grass subfamilies have been analyzed (Table 2). The rice subspecies japonica was sequenced to high accuracy using Sanger sequencing of physically mapped BACs [16] and the indica subspecies was Sanger sequenced using a whole genome shotgun strategy [17]. The relatively compact 389 Mb rice genomes were assembled reasonably well and continued manual annotation of rice genes [18•] formed the RAP2 canonical set that has proved to be instrumental for gene analysis in all grasses. More recently the sorghum genome was Sanger sequenced using a whole genome shotgun approach [19••]. The sequence reads were assembled into >3000 sequence scaffolds, 127 of which were assembled into pseudomolecules containing nearly 90% of the sequence generated. The total genome size including estimated gaps is 758 Mb. Thus the 626 Mb contained in the genome assemblies represents 83% of the estimated genome. The larger genome size of sorghum compared to rice is mainly owing to expanded LTR retroelement populations, such that the euchromatic component of both genomes is between 250 and 300 Mb. Genome sequence and assembly of diploid *B. distachyon*, a wild pooid grass, revealed a compact genome size of 272 Mb and a relatively low repeat content [13••]. Partly due to this low repeat content, the sequence assembly was of unprecedented quality for a draft plant genome with 99.6% of the sequence included in the five pseudomolecules and predicted gaps amounting to only 0.4%. A total of 25 532 protein-coding gene loci was predicted, and over 90% of the predicted coding sequences were supported by Illumina RNA-seq data. Recently the maize B73 genome was sequenced using Sanger sequencing of physically mapped BACs [20••]. Though the abundance of repetitive DNA prevented unambiguous assembly and gene ordering on most BACs, the 2.0 Gb assembly, generated from BAC sequences, was predicted to contain 32 540 protein-coding genes, providing a reasonably narrow range of haploid gene content from a broad diversity of grasses between 25 000 and 35 000 (Table 2).

Comparisons of the complete genomes of three grass species representing a broad diversity of grasses (rice, sorghum and Brachypodium) have identified five major drivers of gene diversification: gene loss following whole genome duplication; tandem duplication; recombination; DNA transposon-mediated exon shuffling; and genome expansion. Grasses are derived from a common ancestor that is predicted to have undergone whole genome duplication between 60 and 70 mya (million years ago) [21•]. The lineages leading to maize underwent further whole genome duplication between 5 and 12 mya, while the bread wheat genome is an allohexaploid derived from three different ancestral genomes. Comparison of rice and sorghum suggests that loss of duplicate genes predated their divergence, resulting in duplicate blocks with approximately 70% of genes having lost their duplicates arising from the pan-grass duplication [19••]. Tandem gene duplicates represent approximately 13% of grass genes, and among grass-specific genes with enriched GO categories tandem duplicates represent nearly 27% of genes [13••]. This demonstrates the importance of tandem duplications in forming new grass-specific gene functions. Comparison of evolutionary changes between small-sized and large-sized grass genomes (Ref. [22•] showed that much higher rates of change were associated with genome expansion compared to changes between subfamilies). More frequent retroelement insertion and inter-element recombination associated with genome expansion is also known to disrupt genes and gene islands [23], leading to increased rates of gene change. Recombination plays a major role in maintaining syntenic gene order by removing retroelements that tend to disrupt gene order, leading to gene-rich distal regions of grass

Table 2

Comparison of haploid protein-coding gene numbers in sequenced grass genomes

	Maize B73	Sorghum	Rice	Brachypodium	Arabidopsis
Version	ZmB73v1	V1.4	RAP2	V1.0	TAIR 8
Genome size (Mb)	2048	758	382	272	119
Non-repeat size (Mb)	425	309	252	200	110
Protein-coding gene loci	32 450	27 640	28 236	25 532	26 990

chromosomes with relatively high conserved gene order [24]. Not all classes of genes behave similarly during grass genome evolution. The largest and most divergent gene families, such as those encoding F-box proteins and NBS-LLR disease resistance proteins, are almost never found in syntenic order, consistent with selection pressures that maintain diversity generated by duplication [13••].

There is an urgent need to obtain useful and comprehensive genome sequence and to access genetic diversity in currently un-sequenced food crops such as wheat, barley and sugarcane, and biomass crops such as *Miscanthus giganteus* and switchgrass (*Panicum virgatum*), in order to accelerate development of these crops for sustainable food and fuel production. However, their size and polyploid complexity (bread wheat is hexaploid and cultivated *Miscanthus* is triploid) are large barriers. The knowledge now available from comparing a broad diversity of grass genome sequences provides a robust framework for interpreting these more complex grass genomes. For example, the exceptionally large genome sizes of barley (5.1 Gb) and wheat (17 Gb) represent major challenges to current genome analysis methods. Syntenic gene order between rice and wheat was used as one of several strategies in the *tour de force* construction of a physical map of BACs representing the 995 Mb wheat chromosome 3B [25••]. BACs in the minimal tiling path can now be sequenced to generate an accurate sequence of genes in their correct chromosome locations. In a complementary approach, an integrated sequence of the 622 Mb barley chromosome 1H was constructed by synteny-based analysis of low coverage shotgun sequence [26•].

Next generation sequencing has the capacity to generate reference sequences and to assay genetic diversity in multiple genotypes [27]. For example, Illumina sequencing provides cost-effective coverage for even the 17 Gb wheat genome, and a range of assembly algorithms [28–30] show promise for assembling low-copy genic sequence. When used on wheat chromosome arms purified from nullisomic lines [31], each of which is about 500–600 Mb, the scale of the assembly problem is significantly reduced and homoeologous relationships are identified. Intergenic regions in wheat and barley are approximately 100 kb and are comprised mainly of nested retroelement insertions. It is unlikely that these regions can be assembled from chromosome-derived short reads using current assembly methods. Therefore larger-scale order available from physically mapped BACs can be used, for example by sequencing the minimal tiling path as pools of BACs. In either case, it is both unnecessary to sequence intergenic sequences of larger grass genomes (and it is currently unlikely that current next generation sequencing and assembly methods would be able to generate sufficiently long assemblies). Instead the approximate order of accurately and completely sequenced genes defined by synteny provides a very useful starting point for map-based gene isolation, assessing diversity of complete gene sets from multiple lines, the dissection of complex traits as QTLs, and molecular breeding.

Developing *B. distachyon* as a model system for grass research

Funding agencies, notably the U.S. Department of Energy, have invested significantly in the development of Brachypodium as a model system to allow researchers to learn the genetic mechanisms controlling traits such as cell wall composition, biomass yield, stress tolerance, and other phenotypes relevant to biomass crop development [5]. This knowledge will be used to accelerate the domestication of wild grasses (e.g. switchgrass and *Miscanthus*) that are promising biomass crops. Comparative analysis of three diverse grass genomes [13••] showed very similar gene contents and gene family composition, supporting the use of Brachypodium as such a general model for diverse grasses. Of particular interest in the context of biomass crops, Brachypodium has a typical grass cell wall structure (Type II). Grass cell walls differ substantially from the Type I cell walls found in dicots, including Arabidopsis. Major differences include the type of hemicellulose (primarily xyloglucans in dicots and glucuronoarabinoxylans in grasses), the presence of high levels of pectin and proteins in dicots, the presence of cross-linking phenolic compounds and mixed linkage glucans in grasses [32]. Further supporting Brachypodium as model for grass cell walls is the demonstration that Brachypodium and *Miscanthus* have similar cell wall compositions [33] and that members of cell wall biosynthetic enzyme families are very similar between Brachypodium, rice, and sorghum [13]. A comparison of salient information about grass models and crops is shown in Table 2.

Practical aspects related to growing Brachypodium have developed concurrently with the emergence of genome resources [34]. One significant advance was the discovery that long days eliminated the need for vernalization in some diploid lines [34]. Modulating day length permits rapid generation turnover when needed or development of larger plants to obtain larger amounts of seeds. For instance, under 24 h days, *B. distachyon* variety Bd21 has been observed to develop from seed to floral spike emergence in 17 days [35]. Similarly, efficient methods for growth of winter habit diploids are also now well established. One such method involves simply planting seeds into wet soil medium, covering with clear plastic wrap to prevent desiccation, and leaving the pots in a cold room for several weeks. Seedlings will emerge and be vernalized during this period, and thus when removed from the cold they will rapidly progress to flowering. Given the small size of Brachypodium anthers and the overall small structure of the inflorescence, crossing initially proved to be a challenge. However, successful

crosses have been made in several labs [35], resulting in many F_2 populations, one of which was subsequently used for development of the first genetic linkage map of the species [36•] and the first recombinant inbred populations (D.F. Garvin, unpublished). Additional efforts have dramatically improved crossing success rates, and detailed instructional guides for routine crossing are available (http://www.ars.usda.gov/SP2UserFiles/person/1931/BrachypodiumCrossing.pdf; http://brachypodium.pw.usda.gov/).

An efficient transformation system is an absolute requirement for a modern model system. Thus, the development of highly efficient methods for *Agrobacterium*-mediated transformation of diploid Brachypodium lines [37•,38•] has been key in establishing Brachypodium as a model system. Significant improvements to published protocols have recently been achieved such that transformation efficiencies of 50% are routine in a production setting (see http://brachypodium.pw.usda.gov/for up-to-date protocols). Thus Brachypodium transformation is at least as efficient as rice transformation [39] making the creation of a population of sequence indexed insertional mutants by T-DNA tagging feasible. Two groups have begun to build such a collection and >10 000 lines are available as of the time of writing (http://brachypodium.pw.usda.gov/TDNA/ and http://www.brachytag.org/), and methods and vectors for efficient T-DNA mutagenesis and sequencing of flanking DNA are now in place to allow other groups to efficiently contribute to making large populations of T-DNA insertion lines for reverse genetics. Chemical and radiation mutageneses provide different spectra of mutants and are commonly used for forward genetic screens, and TILLING populations are increasingly used for reverse genetics in grasses [40]. Methods for chemical mutagenesis by ethyl methanesulfonate (http://brachypodium.pw.usda.gov) and fast neutron (D. Laudencia-Chingcuanco and M. Byrne, pers. com.) mutagenesis of Brachypodium have been established. Finally, the successful demonstration of Virus Induced Gene Silencing (VIGS) in Brachypodium adds another functional genomics resource to the Brachypodium toolbox [41].

As an undomesticated grass, Brachypodium possesses considerable natural diversity that can be used to understand gene function. Three distinct chromosome numbers (n = 5, 10, 15) have been reported for Brachypodium [11,42,43]. Recent cytogenetic evidence suggests that this is not a simple polyploid series as was initially thought; rather, both the 1n = 5 and 1n = 10 cytotypes appear to be true diploids and the 1n = 15 seems to be a tetraploid derived from progenitor genomes similar to the 1n = 5 and 1n = 10 cytotypes [42,43]. Thus, the species *B. distachyon* should probably be considered three distinct species. For use as a model system, the 1n = 5 cytotype is of primary interest. Historical collections dating back to the 1940s are available through the USDA National Plant Germplasm System (NPGS) (www.ars-grin.gov/npgs/). These collections have been characterized and inbred lines developed and made freely available for several diploid accessions [35,38•]. Another collection derived both from NPGS material and independent collections is maintained at the University of Wales, Aberystwyth [11], and is available via a material transfer agreement. The recent introduction of 188 freely available diploid inbred lines generated from seeds collected at 53 sites spread across Turkey has greatly expanded the germplasm available to researchers [44,45•]. Considerable genetic and phenotypic (seed size, flowering time, growth habit, etc.) diversity was observed in this collection. Additional collections have recently been made in Spain that promise to provide additional germplasm in the future (Luis Mur, pers. comm.). To facilitate the study of natural diversity, a project to re-sequence six diverse accessions is underway through the DOE JGI Community Sequencing Program (http://www.jgi.doe.gov/sequencing/why/Bdistachyon.html).

Future direction

With the completion of the Brachypodium genome sequencing project, a key foundation is now in place to allow it to be used as a modern model system. Perhaps just as importantly, a rapidly growing number of researchers have adopted Brachypodium as a model for their research programs. As a measure of its growing acceptance, we have dispatched seed to over 350 labs worldwide and numerous secondary distributions are presumed to have occurred. When combined with growing investments by funding agencies worldwide in research focused on developing improved grass crops for food an fuel, it seems likely that Brachypodium will continue its upward trajectory to join the model organism club. Many areas of plant biology will directly benefit from the use of Brachypodium as a model system, shown by reports covering a wide range of topics (e.g. vernalization and flowering time [46–48], seed storage proteins [49–52], fatty acid turnover [53], plant–pathogen interactions [54], and wounding/insect responses [55,56]). Root biology is another area in which Brachypodium will be useful because, unlike Arabidopsis, it readily forms mycorrhizal associations (M. Harrison and M. Watt, pers. comm.). These associations are important for many crops, especially for the uptake of P, and are especially important for low-input agriculture as envisioned for the future production of biomass crops. Brachypodium root development is very similar to wheat [57], further indicating that Brachypodium will be useful as a model in this area. We have also noted that Brachypodium is very sensitive to pathogens [58] and toxins in the soil. Thus Brachypodium research is poised to make key contributions to understanding grass biology and to sustainable food and fuel production.

Acknowledgements

J.V. was supported by USDA CRIS project 5325-21000-013-00 and by the Office of Science (BER), U.S. Department of Energy, Interagency Agreement No. DE-AI02-07ER64452. M.W.B. is supported by grants from the BBSRC, the EC and the John Innes Centre. DG is supported by USDA ARS CRIS project 3640-21000-021-00D.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Watson L, Dallwitz MJ: In *The Grass Genera of the World.* Edited by Wallingford CT. CAB International; 1992.

2. FAO: *World Agriculture: towards 2030/2050. Interim Report*. Rome, Italy: Global Perspective Studies Unit, Food and Agriculture Organisation of the United Nations; 2006.

3. US Census Bureau Report; 2009.

4. Somerville C: **The billion-ton biofuels vision**. *Science* 2006, **312**:1277.

5. DOE (Ed). *Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda*: U.S. Department of Energy, Office of Science and Office of Energy Efficiency; 2006, available online at http://genomicsgtl.energy.gov/biofuels/b2bworkshop.shtml.

6. Tilman D, Socolow R, Foley JA, Hill J, Larson E, Lynd L, Pacala S, Reilly J, Searchinger T, Somerville C *et al.*: **Energy. Beneficial biofuels—the food, energy, and environment trilemma**. *Science* 2009, **325**:270-271.

7. Bevan M, Walsh S: **The Arabidopsis genome: a foundation for plant research**. *Genome Res* 2005, **15**:1632-1642.

8. Kellogg EA: **Evolutionary history of the grasses**. *Plant Physiol* 2001, **125**:1198-1205.

9. Moore G, Devos KM, Wang Z, Gale MD: **Cereal genome evolution. Grasses, line up and form a circle**. *Curr Biol* 1995, **5**:737-739.

10. Bablak P, Draper J, Davey MR, Lynch PT: **Plant regeneration and micropropagation of *Brachypodium distachyon***. *Plant Cell Tissue Org Cult* 1995, **42**:97-107.

11. Draper J, Mur LA, Jenkins G, Ghosh-Biswas GC, Bablak P, Hasterok R, Routledge AP: ***Brachypodium distachyon*. A new model system for functional genomics in grasses**. *Plant Physiol* 2001, **127**:1539-1555.

12. Garvin DF, Gu Y-Q, Hasterok R, Hazen SP, Jenkins G, Mockler TC, Mur LAJ, Vogel JP: **Development of genetic and genomic research resources for *Brachypodium distachyon*, a new model system for grass crop research**. *Plant Genome* 2008, **48**:69-84.

13. •• International Brachypodium Initiative: **Genome sequencing and analysis of the model grass *Brachypodium distachyon***. *Nature* 2010 doi: 10.1038/nature08747.
This paper describes the sequencing, annotation and comparative analysis of the Brachypodium genome. This work was notable for its unprecedented completeness of the assembly, the accuracy of the annotation and the comparison, for the first time, of whole genomes from representatives of the three most economically important grass lineages.

14. Wicker T, Keller B: **Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families**. *Genome Res* 2007, **17**:1072-1081.

15. Davies TJ, Barraclough TG, Chase MW, Soltis PS, Soltis DE, Savolainen V: **Darwin's abominable mystery: insights from a supertree of the angiosperms**. *Proc Natl Acad Sci U S A* 2004, **101**:1904-1909.

16. International Rice Genome Sequencing Project: **The map-based sequence of the rice genome**. *Nature* 2005, **436**:793-800.

17. Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X *et al.*: **A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica)**. *Science* 2002, **296**:79-92.

18. • Tanaka T, Antonio BA, Kikuchi S, Matsumoto T, Nagamura Y, Numa H, Sakai H, Wu J, Itoh T, Sasaki T *et al.*: **The rice annotation project database (RAP-DB): 2008 update**. *Nucleic Acids Res* 2008, **36**:D1028-1033.
This paper describes a comprehensive manual annotation of the rice genome that provides a 'gold standard' for annotating genes from other grass genomes.

19. •• Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A *et al.*: **The *Sorghum bicolor* genome and the diversification of grasses**. *Nature* 2009, **457**:551-556.
The sequence and analysis of the Sorghum genome described here is a major achievement in plant genome research because it is the first panicoid grass sequenced and because it is one of the largest genomes that is essentially completely assembled using whole genome shotgun sequences.

20. •• Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA *et al.*: **The B73 maize genome: complexity, diversity, and dynamics**. *Science* 2009, **326**:1112-1115.
The sequence of the maize genome is a major milestone in plant biology due to the size and complexity of its genome and the great significance of the plant as one of the world's most important food and fuel crops, and because of its central importance in plant biology.

21. • Salse J, Bolot S, Throude M, Jouffe V, Piegu B, Quraishi UM, Calcagno T, Cooke R, Delseny M, Feuillet C: **Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution**. *Plant Cell* 2008, **20**:11-24.
This paper describes a comprehensive bioinformatic analysis of syntenic relationships among the grasses and provides a framework for understanding genome evolution.

22. • Luo MC, Deal KR, Akhunov ED, Akhunova AR, Anderson OD, Anderson JA, Blake N, Clegg MT, Coleman-Derr D, Conley EJ *et al.*: **Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae**. *Proc Natl Acad Sci U S A* 2009, **106**:15780-15785.
The insertion of grass chromosomes into each other was shown to be an important and distinctive mechanism of chromosome evolution in the grasses.

23. Vicient CM, Kalendar R, Schulman AH: **Variability, recombination, and mosaic evolution of the barley BARE-1 retrotransposon**. *J Mol Evol* 2005, **61**:275-291.

24. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW: **Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis**. *Plant Cell* 2003, **15**:809-834.

25. •• Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeyer W *et al.*: **A physical map of the 1-gigabase bread wheat chromosome 3B**. *Science* 2008, **322**:101-104.
Constructing the first physical map of a wheat chromosome was a *tour de force* of physical mapping and establishes the feasibility of making maps of all wheat chromosomes to facilitate gene isolation and possibly to sequence the wheat genome.

26. • Mayer KF, Taudien S, Martis M, Simkova H, Suchankova P, Gundlach H, Wicker T, Petzold A, Felder M, Steuernagel B *et al.*: **Gene content and virtual gene order of barley chromosome 1H**. *Plant Physiol* 2009, **151**:496-505.
The concept of a 'syntenic build' as a first step towards constructing useful sequence assemblies of large grass chromosomes is demonstrated in this important publication.

27. Nordborg M, Weigel D: **Next-generation genetics in plants**. *Nature* 2008, **456**:720-723.

28. Zerbino DR, Birney E: **Velvet: algorithms for de novo short read assembly using de Bruijn graphs**. *Genome Res* 2008, **18**:821-829.

29. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I: **ABySS: a parallel assembler for short read sequence data**. *Genome Res* 2009, **19**:1117-1123.

30. Zerbino DR, McEwen GK, Margulies EH, Birney E: **Pebble and rock band: heuristic resolution of repeats and scaffolding in the velvet short-read de novo assembler**. *PLoS One* 2009, **4**:e8407.

31. Dolezel J, Kubalakova M, Paux E, Bartos J, Feuillet C: **Chromosome-based genomics in the cereals**. *Chromosome Res* 2007, **15**:51-66.

32. Vogel J: **Unique aspects of the grass cell wall**. *Curr Opin Plant Biol* 2008, **11**:301-307.

33. Gomez LD, Bristow JK, Statham ER, McQueen-Mason SJ: **Analysis of saccharification in *Brachypodium distachyon* stems under mild conditions of hydrolysis**. *Biotechnol Biofuels* 2008, **1**:15.

34. Vogel JP, Garvin DF, Leong OM, Hayden DM: **Agrobacterium-mediated transformation and inbred line development in the model grass *Brachypodium distachyon***. *Plant Cell Tissue Org Cult* 2006, **84**:199-211.

35. Garvin DF: **Brachypodium distachyon: a new model system for structural and functional analysis of grass genomes**. In *Model plants and crop improvement.* Edited by Varshney RK, Koebner RMD. Boca Raton, Fla.: Taylor and Francis; 2007: 109-123.

36. • Garvin DF, McKenzie N, Vogel JP, Mockler TC, Blankenheim ZJ, Wright J, Cheema JJS, Dicks J, Huo N, Hayden DM *et al.*: **An SSR-based genetic linkage map of the model grass *Brachypodium distachyon***. *Genome* 2010, **53**:1-13.
This paper describes the first genetic map of *Brachypodium distachyon,* a key initial milestone in genome analysis and establishing it as a biological system.

37. • Vain P, Worland B, Thole V, McKenzie N, Alves SC, Opanowicz M, Fish LJ, Bevan MW, Snape JW: **Agrobacterium-mediated transformation of the temperate grass *Brachypodium distachyon* (genotype Bd21) for T-DNA insertional mutagenesis**. *Plant Biotechnol J* 2008, **6**:236-245.
This paper and the following paper demonstrated highly efficient *Agrobacterium*-mediated transformation of Brachypodium. These papers clearly placed Brachypodium in the elite group of grasses that posses a transformation system efficient enough for a modern model system. In addition, this paper presented optimized methods for the sequencing of DNA flanking insertion sites.

38. • Vogel J, Hill T: **High-efficiency Agrobacterium-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3**. *Plant Cell Rep* 2008, **27**:471-478.
This paper is a follow up to this group's initial demonstration of *Agrobacterium*-mediated transformation [32] of Brachypodium. In addition to improving transformation efficiency 10-fold, they created an inbred line that was more efficiently transformed.

39. Tyagi AK, Mohanty A: **Rice transformation for crop improvement and functional genomics**. *Plant Sci* 2000, **158**:1-18.

40. Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S, Comai L: **Discovery of chemically induced mutations in rice by TILLING**. *BMC Plant Biol* 2007, **7**:19.

41. Demircan T, Akkaya MS: **Virus-induced gene silencing in *Brachypodium distachyon*, a model organism for cereals**. *Plant Cell Tissue Org Cult* 2010, **100**:91-96.

42. Hasterok R, Draper J, Jenkins G: **Laying the cytotaxonomic foundations of a new model grass. *Brachypodium distachyon* (L.) Beauv**. *Chromosome Res* 2004, **12**:397-403.

43. Hasterok R, Marasek A, Donnison IS, Armstead I, Thomas A, King IP, Wolny E, Idziak D, Draper J, Jenkins G: **Alignment of the genomes of *Brachypodium distachyon* and temperate cereals and grasses using bacterial artificial chromosome landing with fluorescence in situ hybridization**. *Genetics* 2006, **173**:349-362.

44. Filiz E, Ozdemir BS, Budak F, Vogel JP, Tuna M, Budak H: **Molecular, morphological, and cytological analysis of diverse *Brachypodium distachyon* inbred lines**. *Genome* 2009, **52**:876-890.

45. • Vogel JP, Tuna M, Budak H, Huo N, Gu YQ, Steinwand MA: **Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon***. *BMC Plant Biol* 2009, **9**:88.
This paper describes an extensive collection of natural accessions in terms of phenotypic differences and genotypic differences based on SSR markers. The authors convincingly demonstrate that there is sufficient diversity in the collection to allow Brachypodium to be used as a model for natural variation and to allow efficient map-based cloning approaches.

46. Faricelli ME, Valarik M, Dubcovsky J: **Control of flowering time and spike development in cereals: the earliness per se Eps-1 region in wheat, rice, and Brachypodium**. *Funct Integr Genomics* 2009 doi: 10.1007/s10142-009-0146-7.

47. Olsen P, Lenk I, Jensen CS, Petersen K, Andersen CH, Didion T, Neilsen KK: **Analysis of two heterologus flowering genes in *Brachypodium distachyon* demonstrates its potential as a grass model plant**. *Plant Sci* 2006, **170**:1020-1025.

48. Schwartz CJ, Doyle MR, Manzaneda AJ, Rey PJ, Mitchell-Olds T, Amasino RM: **Natural variation of flowering time and vernalization responsiveness in Brachypodium distachyon**. *Bioenergy Res* 2010, **3**:38.

49. Charles M, Tang H, Belcram H, Paterson A, Gornicki P, Chalhoub B: **Sixty million years in evolution of soft grain trait in grasses: emergence of the softness locus in the common ancestor of Pooideae and Ehrhartoideae, after their divergence from Panicoideae**. *Mol Biol Evol* 2009, **26**:1651-1661.

50. Gu YQ, Wanjugi H, Coleman-Derr D, Kong X, Anderson OD: **Conserved globulin gene across eight grass genomes identify fundamental units of the loci encoding seed storage proteins**. *Funct Integr Genomics* 2010, **10**:111-122.

51. Laudencia-Chingcuanco DL, Vensel WH: **Globulins are the main seed storage proteins in *Brachypodium distachyon***. *Theor Appl Genet* 2008, **117**:555-563.

52. Xu JH, Messing J: **Amplification of prolamin storage protein genes in different subfamilies of the Poaceae**. *Theor Appl Genet* 2009, **119**:1397-1412.

53. Yang Z, Ohlrogge JB: **Turnover of fatty acids during natural senescence of Arabidopsis, Brachypodium, and switchgrass and in Arabidopsis beta-oxidation mutants**. *Plant Physiol* 2009, **150**:1981-1989.

54. Parker D, Beckmann M, Enot DP, Overy DP, Rios ZC, Gilbert M, Talbot N, Draper J: **Rice blast infection of *Brachypodium distachyon* as a model system to study dynamic host/pathogen interactions**. *Nat Protoc* 2008, **3**:435-445.

55. Azhaguvel P, Lia W, Rudda JC, Gill BS, Michels GJ, Weng Y: **Aphid feeding response and microsatellite-based genetic diversity among diploid *Brachypodium distachyon* accessions**. *Plant Genet Resour* 2009, **7**:72-79.

56. Mur LAJ, Xu R, Casson SA, Stoddart WM, Routledge APM, Draper J: **Characterisation of a proteinase inhibitor from *Brachypodium distachyon* suggests the conservation of defence signalling pathways between grasses and dicots**. *Mol Plant Pathol* 2004, **5**:267-280.

57. 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 crops**. *Funct Plant Biol* 2009, **36**:960.

58. Vogel JP, Bragg JN: ***Brachypodium distachyon*, a New Model for the Triticeae**. In *Genetics and Genomics of the Triticeae.* Edited by Feuillet C, Muehlbauer G. Springer; 2009:427-449. . [Jorgensen RA (Series Editor): Plant Genetics and Genomics: Crops and Models, vol 7.]