Grasses provide the bulk of human nutrition, and highly productive grasses are promising sources of sustainable energy. The grass family (Poaceae) comprises over 600 genera and more than 10,000 species that dominate many ecological and agricultural systems. So far, genomic efforts have largely focused on two economically important grass subfamilies, the Ehrhartoideae (rice) and the Panicoideae (maize, sorghum, sugarcane and millets). The rice and sorghum genome sequences and a detailed physical map of maize showed extensive conservation of gene order and both ancient and relatively recent polyploidization.

Most cool season cereal, forage and turf grasses belong to the Pooideae subfamily, which is also the largest grass subfamily. The genomes of many pooids are characterized by daunting size and complexity. For example, the bread wheat genome is approximately 17,000 megabases (Mb) and contains three independent genomes. This has prohibited genome-scale comparisons spanning the three most economically important grass subfamilies.

**Brachypodium**, a member of the Pooideae subfamily, is a wild annual grass endemic to the Mediterranean and Middle East that has promise as a model system. This has led to the development of highly efficient transformation, germplasm collections, genetic markers, a genetic linkage map, bacterial artificial chromosome (BAC) libraries, physical maps (M.F., unpublished observations), mutant collections (http://brachypodium.pw.usda.gov, http://www.brachytag.org), microarrays and databases (http://www.brachybase.org, http://www.phytozome.net, http://www.modelcrop.org, http://mips.helmholtz-muenchen.de/plant/index.jsp) that are facilitating the use of Brachypodium by the research community. The genome sequence described here will allow Brachypodium to act as a powerful functional genomics resource for the grasses. It is also an important advance in grass structural genomics, permitting, for the first time, whole-genome comparisons between members of the three most economically important grass subfamilies.

**Genome sequence assembly and annotation**

The diploid inbred line Bd21 (ref. 19) was sequenced using whole-genome shotgun sequencing (Supplementary Table 1). The ten largest scaffolds contained 99.6% of all sequenced nucleotides (Supplementary Table 2). Comparison of these ten scaffolds with a genetic map (Supplementary Fig. 1) detected two false joins and created a further seven joins to produce five pseudomolecules that spanned 272 Mb (Supplementary Table 3), within the range measured by flow cytometry. The assembly was confirmed by cytogenetic analysis (Supplementary Fig. 2) and alignment with two physical maps and sequenced BACs (Supplementary Data). More than 98% of expressed sequence tags (ESTs) mapped to the sequence assembly, consistent with a near-complete genome (Supplementary Table 4 and Supplementary Fig. 3). Compared to other grasses, the Brachypodium genome is very compact, with retrotransposons concentrated at the centromeres and syntenic breakpoints (Fig. 1). DNA transposons and derivatives are broadly distributed and primarily associated with gene-rich regions.

We analysed small RNA populations from inflorescence tissues with deep Illumina sequencing, and mapped them onto the genome sequence (Fig. 2a, Supplementary Fig. 4 and Supplementary Table 5). Small RNA reads were most dense in regions of high repeat density, similar to the distribution reported in Arabidopsis. We identified 413 and 198 21- and 24-nucleotide phased short interfering RNA (siRNA) loci, respectively. Using the same algorithm, the only phased loci identified in Arabidopsis were five of the eight trans-acting siRNA loci, and none was 24-nucleotide phased. The biological functions of these clusters of Brachypodium phased siRNAs, which account for a significant number of small RNAs that map outside repeat regions, are not known at present.

A total of 25,532 protein-coding gene loci was predicted in the v1.0 annotation (Supplementary Information and Supplementary Table 6). This is in the same range as rice (RAP2, 28,236) and sorghum (v1.4, 27,640), suggesting similar gene numbers across a broad diversity of grasses. Gene models were evaluated using ~10.2 gigabases (Gb) of Illumina RNA-seq data (Supplementary Fig. 5). Overall, 92.7% of predicted coding sequences (CDS) were supported by Illumina RNA-seq data (Supplementary Fig. 2b), demonstrating the high accuracy of the Brachypodium gene predictions. These gene models are available from several databases (such as http://www.brachybase.org, http://www.phytozome.net, http://www.modelcrop.org and http://mips.org).

Between 77 and 84% of gene families (defined according to Supplementary Fig. 6) are shared among the three grass subfamilies represented by Brachypodium, rice and sorghum, reflecting a relatively
The abundance and distribution of the following genome elements are shown: complete LTR retroelements (cLTRs); solo-LTRs (sLTRs); potentially autonomous DNA transposons that are not miniature inverted-repeat transposable elements (MITEs) (DNA-TEs); MITEs; gene exons (CDS); gene introns and satellite tandem arrays (STA). Graphs are from 0 to 100 percent base-pair (%bp) coverage of the respective window. The heat map tracks have different ranges and different maximum (max) pseudocolor levels: STA (0–55, scaled to max 100) %bp; cLTRs (0–36, scaled to max 20) %bp; sLTRs (0–4) %bp; DNA-TEs (0–20) %bp; MITEs (0–22) %bp; CDS (exons) (0–22.3) %bp. The triangles identify syntenic breakpoints.

recent common origin (Fig. 2c). Grass-specific genes include transmembrane receptor protein kinases, glycosyltransferases, peroxidases and P450 proteins (Supplementary Table 7B). The Pooidae-specific gene set contains only 265 gene families (Supplementary Table 7C) comprising 811 genes (1,400 including singletons). Genes enriched in grasses were significantly more likely to be contained in tandem arrays than random genes, demonstrating a prominent role for tandem gene expansion in the evolution of grass-specific genes (Supplementary Fig. 7 and Supplementary Table 8).

To validate and improve the v1.0 gene models, we manually annotated 2,755 gene models from 97 diverse gene families (Supplementary Tables 9–11) relevant to bioenergy and food crop improvement. We annotated 866 genes involved in cell wall biosynthesis/modification and 948 transcription factors from 16 families39. Only 13% of the gene models required modification and very few pseudogenes were identified, demonstrating the accuracy of the v1.0 annotation. Phylogenetic trees for 62 gene families were constructed using genes from rice, Arabidopsis, sorghum and poplar. In nearly all cases, Brachypodium genes had a similar distribution to rice and sorghum, demonstrating that Brachypodium is suitably generic for grass functional genomics research (Supplementary Figs 8 and 9). Analysis of the predicted secretome identified substantial differences in the distribution of cell wall metabolism genes between dicots and grasses (Supplementary Tables 12, 13 and Supplementary Fig. 10), consistent with their different cell walls26. Signal peptide probability curves also suggested that start codons were accurately predicted (Supplementary Fig. 11).

Maintaining a small grass genome size

Exhaustive analysis of transposable elements (Supplementary Information and Supplementary Table 14) showed retrotransposon sequences comprise 21.4% of the genome, compared to 26% in rice, Brachypodium, suggests that gene expansion in the evolution of grass-specific genes is much lower than random genes, demonstrating a prominent role for tandem gene expansion in the evolution of grass-specific genes (Supplementary Fig. 7 and Supplementary Table 8).

To validate and improve the v1.0 gene models, we manually annotated 2,755 gene models from 97 diverse gene families (Supplementary Tables 9–11) relevant to bioenergy and food crop improvement. We annotated 866 genes involved in cell wall biosynthesis/modification and 948 transcription factors from 16 families39. Only 13% of the gene models required modification and very few pseudogenes were identified, demonstrating the accuracy of the v1.0 annotation. Phylogenetic trees for 62 gene families were constructed using genes from rice, Arabidopsis, sorghum and poplar. In nearly all cases, Brachypodium genes had a similar distribution to rice and sorghum, demonstrating that Brachypodium is suitably generic for grass functional genomics research (Supplementary Figs 8 and 9). Analysis of the predicted secretome identified substantial differences in the distribution of cell wall metabolism genes between dicots and grasses (Supplementary Tables 12, 13 and Supplementary Fig. 10), consistent with their different cell walls26. Signal peptide probability curves also suggested that start codons were accurately predicted (Supplementary Fig. 11).

Maintaining a small grass genome size

Exhaustive analysis of transposable elements (Supplementary Information and Supplementary Table 14) showed retrotransposon sequences comprise 21.4% of the genome, compared to 26% in rice, Brachypodium, suggesting that gene expansion in the evolution of grass-specific genes is much lower than random genes, demonstrating a prominent role for tandem gene expansion in the evolution of grass-specific genes (Supplementary Fig. 7 and Supplementary Table 8).

To validate and improve the v1.0 gene models, we manually annotated 2,755 gene models from 97 diverse gene families (Supplementary Tables 9–11) relevant to bioenergy and food crop improvement. We annotated 866 genes involved in cell wall biosynthesis/modification and 948 transcription factors from 16 families39. Only 13% of the gene models required modification and very few pseudogenes were identified, demonstrating the accuracy of the v1.0 annotation. Phylogenetic trees for 62 gene families were constructed using genes from rice, Arabidopsis, sorghum and poplar. In nearly all cases, Brachypodium genes had a similar distribution to rice and sorghum, demonstrating that Brachypodium is suitably generic for grass functional genomics research (Supplementary Figs 8 and 9). Analysis of the predicted secretome identified substantial differences in the distribution of cell wall metabolism genes between dicots and grasses (Supplementary Tables 12, 13 and Supplementary Fig. 10), consistent with their different cell walls26. Signal peptide probability curves also suggested that start codons were accurately predicted (Supplementary Fig. 11).
54% in sorghum, and more than 80% in wheat. Thirteen retroelement sets were younger than 20,000 years, showing a recent activation compared to rice (Supplementary Fig. 12), and a further 53 retroelement sets were less than 0.1 million years (Myr) old. A minimum of 17.4 Mb has been lost by long terminal repeat (LTR)–LTR recombination, demonstrating that retroelement expansion is countered by removal through recombination. In contrast, retroelements persist for very long periods of time in the closely related Triticeae.

DNA transposons comprise 4.77% of the Brachypodium genome, within the range found in other grass genomes. Transcriptome data and structural analysis suggest that many non-autonomous Mariner DTT and Harbinger elements recruit transposases from other families. Two CACTA DTC families (M and N) carried five non-element genes, and the Harbinger U family has amplified a NBS-LRR gene family (Supplementary Figs 13 and 14), adding it to the group of transposable elements implicated in gene mobility. Centromeric regions were characterized by low gene density, characteristic repeats and retroelement clusters (Supplementary Fig. 15). Other repeat classes are described in Supplementary Table 15. Conserved non-coding sequences are described in Supplementary Fig. 16.

**Whole-genome comparison of three diverse grass genomes**

The evolutionary relationships between Brachypodium, sorghum, rice and wheat were assessed by measuring the mean synonymous substitution rates (Ks) of orthologous gene pairs (Supplementary Information, Supplementary Fig. 17 and Supplementary Table 16), from which divergence times of Brachypodium from wheat 32–39 Myr ago, rice 40–53 Myr ago, and sorghum 45–60 Myr ago (Fig. 3a) were estimated. The Ks of orthologous gene pairs in the intragenomic Brachypodium duplications (Fig. 3b) suggests duplication 56–72 Myr ago, before the diversification of the grasses. This is consistent with previous evolutionary histories inferred from a small number of genes.

Paralogous relationships among Brachypodium chromosomes showed six major chromosomal duplications covering 92.1% of the genome (Fig. 3b), representing ancestral whole-genome duplication. Using the rice and sorghum genome sequences, genetic maps of barley and Aegilops tauschii (the D genome donor of hexaploid wheat), and bin-mapped wheat ESTs, 21,045 orthologous relationships between Brachypodium, rice, sorghum and Triticeae were identified (Supplementary Information). These identified 59 blocks of collinear genes covering 99.2% of the Brachypodium genome (Fig. 3c–e). The orthologous relationships are consistent with an evolutionary model that shaped five Brachypodium chromosomes from a five-chromosome ancestral genome by a 12-chromosome intermediate involving seven major chromosome fusions (Supplementary Fig. 18). These collinear blocks of orthologous genes provide a robust and precise sequence framework for understanding grass genome evolution and aiding the assembly of sequences from other pooid grasses. We identified 14 major syntenic disruptions between Brachypodium and rice/sorghum that can be explained by nested insertions of entire chromosomes into centromeric regions (Fig. 4a, b, c). Similar nested insertions in sorghum and barley (Fig. 4c, d) were also identified. Centromeric repeats and peaks in retroelements at the junctions of chromosome insertions are footprints of these insertion events (Supplementary Fig. 15C and Fig. 1), as is higher gene density at the former distal regions of the inserted chromosomes (Fig. 1). Notably, the reduction in chromosome number in Brachypodium and wheat occurred independently because none of the chromosome fusions are shared by Brachypodium and the Triticeae (Supplementary Fig. 18).

---

**Figure 3 | Brachypodium genome evolution and synteny between grass subfamilies.**

**a.** The distribution maxima of mean synonymous substitution rates (Ks) of Brachypodium, rice, sorghum and wheat orthologous gene pairs (Supplementary Table 16) were used to define the divergence times of these species and the age of interchromosomal duplications in Brachypodium. WGD, whole-genome duplication. The numbers refer to the predicted divergence times measured as Myr ago by the NG or ML methods. **b.** Diagram showing the six major interchromosomal Brachypodium duplications, defined by 723 paralogous relationships, as coloured bands linking the five chromosomes. **c.** Identification of chromosome relationships between the Brachypodium, rice and sorghum genomes. Orthologous relationships between the 25,532 protein-coding Brachypodium genes, 7,216 sorghum orthologues (12 syntenic blocks), and 8,533 rice orthologues (12 syntenic blocks) were defined. Sets of collinear orthologous relationships are represented by a coloured band according to each Brachypodium chromosome (blue, chromosome (chr.) 1; yellow, chr. 2; violet, chr. 3; red, chr. 4; green, chr. 5). The white region in each Brachypodium chromosome represents the centromeric region. **d.** Orthologous gene relationships between Brachypodium and barley and Ae. tauschii were identified using genetically mapped ESTs. 2,516 orthologous relationships defined 12 syntenic blocks. These are shown as coloured bands. **e.** Orthologous gene relationships between Brachypodium and hexaploid wheat defined by 5,003 ESTs mapped to wheat deletion bins. Each set of orthologous relationships is represented by a band that is evenly spread across each deletion interval on the wheat chromosomes.
and complex genomes of wheat and barley, two other pooid grasses, demonstrate a substantially higher rate of genome change in {\it Ae. tauschii} (Supplementary Table 17). This may be due to retroelement activity that increases syntenic disruptions, as proposed for chromosome 5S later\textsuperscript{41}. Among seven relatively large gene families, four were highly syntenic and two (NBS-LRR and F-box) were almost never found in syntenic order when compared to rice and sorghum (Supplementary Table 18), typical of the larger genomes of closely related grasses.

Comparisons of evolutionary rates between {\it Brachypodium}, sorghum, rice and {\it Ae. tauschii} demonstrated a substantially higher rate of genome change in {\it Ae. tauschii} (Supplementary Table 17). This may be due to retroelement activity that increases syntenic disruptions, as proposed for chromosome 5S later\textsuperscript{41}. Among seven relatively large gene families, four were highly syntenic and two (NBS-LRR and F-box) were almost never found in syntenic order when compared to rice and sorghum (Supplementary Table 18), consistent with the rapid diversification of the NBS-LRR and F-box gene families\textsuperscript{42}.

![Figure 4](Image) **Figure 4 | A recurring pattern of nested chromosome fusions in grasses.** a. The five {\it Brachypodium} chromosomes are coloured according to homology with rice chromosomes (Os1–Os12). Chromosomes descended from an ancestral chromosome (A4–A11) through whole-genome duplication are shown in shades of the same colour. Gene density is indicated as a red line above the chromosome maps. Major discontinuities in gene density identify syntenic breakpoints, which are marked by a diamond. White diamonds identify fusion points containing remnant centromeric repeats. b. A pattern of nested insertions of whole chromosomes into centromeric regions explains the observed syntenic break points. Bd5 has not undergone chromosome fusion. c. Examples of nested chromosome insertions in sorghum (Sb) chromosomes 1 and 2. d. Examples of nested chromosome insertions in barley (H chromosomes) inferred from genetic maps. Nested insertions were not identified in other chromosomes, possibly owing to the low resolution of genetic maps.

Discussion

As the first genome sequence of a pooid grass, the {\it Brachypodium} genome aids genome analysis and gene identification in the large and complex genomes of wheat and barley, two other pooid grasses that are among the world’s most important crops. The very high quality of the {\it Brachypodium} genome sequence, in combination with those from two other grass subfamilies, enabled reconstruction of chromosome evolution across a broad diversity of grasses. This analysis contributes to our understanding of grass diversification by explaining how the varying chromosome numbers found in the major grass subfamilies derive from an ancestral set of five chromosomes by nested insertions of whole chromosomes into centromeres. The relatively small genome of {\it Brachypodium} contains many active retroelement families, but recombination between these keeps genome expansion in check. The short arm of chromosome 5 deviates from the rest of the genome by exhibiting a trend towards genome expansion through increased retroelement numbers and disruption of gene order more typical of the larger genomes of closely related grasses.

Grass crop improvement for sustainable fuel\textsuperscript{44} and food\textsuperscript{45} production requires a substantial increase in research in species such as {\it Miscanthus}, switchgrass, wheat and cool season forage grasses. These considerations have led to the rapid adoption of {\it Brachypodium} as an experimental system for grass research. The similarities in gene content and gene family structure between {\it Brachypodium}, rice and sorghum support the value of {\it Brachypodium} as a functional genomics model for all grasses. The {\it Brachypodium} genome sequence analysis reported here is therefore an important advance towards securing sustainable supplies of food, feed and fuel from new generations of grass crops.


43. Garvin, D. F. et al. *Brachypodium distachyon* has been deposited at DBT/EMBL/GenBank under the accession ADDNO00000000. (The version described in this manuscript is the first version, accession ADDNO01000000). EST sequences have been deposited with dbEST (accessions 67946317–68053959). We acknowledge the contributions of the late M. Gale, who identified the importance of conserved gene order in grass genomes. This work was supported by the US Department of Energy Joint Genome Institute Community Sequencing Program project with J.P.V., D.F.G., T.C.M. and M.W.B., a BBSRC grant to M.W.B., an EU Contract Agronomics grant to M.W.B. and K.F.X.M., and GABI Barlex grant to K.F.X.M. Illumina transcriptome sequencing was supported by a DOE Plant Feedstock Genomics for Bioenergy grant and an Oregon State Agricultural Research Foundation grant to T.C.M.; small RNA research was supported by the DOE Plant Feedstock Genomics for Bioenergy grants to P.I.G. and T.C.M.; annotation was supported by a DOE Plant Feedstocks for Genomics Bioenergy grant to J.P.V. A full list of support and acknowledgments is in the Supplementary Information.

**Author Information.** The whole-genome shotgun sequence of *Brachypodium distachyon* has been deposited at DDBJ/EMBL/GenBank under the accession ADDNO00000000. (The version described in this manuscript is the first version, accession ADDNO01000000). EST sequences have been deposited with dbEST (accessions 67946317–68053959). We acknowledge the contributions of the late M. Gale, who identified the importance of conserved gene order in grass genomes. This work was supported by the US Department of Energy Joint Genome Institute Community Sequencing Program project with J.P.V., D.F.G., T.C.M. and M.W.B., a BBSRC grant to M.W.B., an EU Contract Agronomics grant to M.W.B. and K.F.X.M., and GABI Barlex grant to K.F.X.M. Illumina transcriptome sequencing was supported by a DOE Plant Feedstock Genomics for Bioenergy grant and an Oregon State Agricultural Research Foundation grant to T.C.M.; small RNA research was supported by the DOE Plant Feedstock Genomics for Bioenergy grants to P.I.G. and T.C.M.; annotation was supported by a DOE Plant Feedstocks for Genomics Bioenergy grant to J.P.V. A full list of support and acknowledgments is in the Supplementary Information. 

**Acknowledgements.**

**Author Contributions.** See list of consortium authors below.

**Funding and Competing Interests.** The International Brachypodium Initiative. Principal investigators: John P. Vogel1, David F. Garvin2, Todd C. Mockler3, Jeremy Schmutz4, Dan Rokhsar5, Michael W. Bevan6,7,8; DNA sequencing and assembly (Curry Barry)2, Susan Lucas2, Miranda Harmon-Smith2, Kathleen Laird3, Hope Tice6, Jeremy Schmutz4 (Leader), Jane Grimwood2, Neil McKenzie6, Michael W. Bevan6; Pseudomolecule assembly and BARC and ending (Naxin Huo1, Yong Q. Gu1, Gerard R. Lazo1, Olin D. Anderson1, John P. Vogel1, Leader, Frank M. You8, Ming-Cheng Luo1, Jan Dvorak2, Jonathan Wright3, Valerie Abernethy5,7, Michael W. Bevan6,7, Dominika Izidzik4, Robertasterok5, David F. Garvin6; Transcriptome sequencing and analysis (Erik Lindquist7, Mei Wang1, Samuel E. Fox1, Henry D. Priest7, Sergei A. Filichkin3, Scott A. Givan8, Douglas W. Bryant9, Jeff H. Chang10, Todd C. Mockler7, Leader, Hayan Wu10,12, Wei Wu10, Qing Hua10, Pin Gu10, Brian Bradzuk2, Todd P. Michael11, Samuel P. Haizen11, Jennifer N. Bragg11, Debbie Lauderdale-Chang12, John T. Roberts12, David F. Garvin12, Xuping Wang12, Neil McKenzie12, Michael W. Bevan6,7, Gene analysis and annotation (George Haba13, Manuel Spannagl12, Klaus Mayer12, Leader, Thomas Rattel12, Therese Mitros12, Dan Rokhsar12, Sankar Jee Lee13, Jocelyn C. K. Rose13, Lukas A. Mueller14, Thomas L. York14, Reprints and permissions information is available online at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.P.V. (john.vogel@ars.usda.gov) or D.F.G. (david.garvin@ars.usda.gov) or T.C.M. (tmockler@cgrb.oregonstate.edu) or M.W.B. (michael.bevan@bbsrc.ac.uk).

**Author Contributions.** See list of consortium authors below.
1USDA-ARS Western Regional Research Center, Albany, California 94710, USA.  
2USDA-ARS Plant Science Research Unit and University of Minnesota, St Paul, Minnesota 55108, USA.  
3Oregon State University, Corvallis, Oregon 97331-4501, USA.  
4HudsonAlpha Institute, Huntsville, Alabama 35806, USA.  
5DOE Joint Genome Institute, Walnut Creek, California 94598, USA.  
6University of California Berkeley, Berkeley, California 94720, USA.  
7John Innes Centre, Norwich NR4 7UJ, UK.  
8University of California Davis, Davis, California 95616, USA.  
9University of Silesia, 40-032 Katowice, Poland.  
10Iowa State University, Ames, Iowa 50011, USA.  
11Washington State University, Pullman, Washington 99163, USA.  
12University of Florida, Gainesville, Florida 32611, USA.  
13Rutgers University, Piscataway, New Jersey 08855-0759, USA.  
14University of Massachusetts, Amherst, Massachusetts 01003-9292, USA.  
15USDA-ARS Vegetable Crops Research Unit, Horticulture Department, University of Wisconsin, Madison, Wisconsin 53706, USA.  
16Helmholtz Zentrum München, D-85764 Neuherberg, Germany.  
17Technical University München, 80333 München, Germany.  
18Cornell University, Ithaca, New York 14853, USA.  
19Boyece Thompson Institute for Plant Research, Ithaca, New York 14853-1801, USA.  
20University of Zurich, 8008 Zurich, Switzerland.  
21MTT Agrifood Research and University of Helsinki, FIN-00014 Helsinki, Finland.  
22Federal University of Pelotas, Pelotas, 96001-970, RS, Brazil.  
23Michigan State University, East Lansing, Michigan 48824, USA.  
24China Agricultural University, Beijing 100094, China.  
25Purdue University, West Lafayette, Indiana 47907, USA.  
26The University of Texas, Arlington, Texas 76019, USA.  
27Institut National de la Recherche Agronomique UMR 1095, 63100 Clermont-Ferrand, France.  
28University of California San Diego, La Jolla, California 92093, USA.  
29National Centre for Genome Resources, Santa Fe, New Mexico 87505, USA.  
30University of Delaware, Newark, Delaware 19716, USA.  
31Joint Bioenergy Institute, Emeryville, California 94720, USA.  
32University of Copenhagen, Frederiksberg DK-1871, Denmark.  
33USDA-ARS Appalachian Fruit Research Station, Kearneysville, West Virginia 25430, USA.  
34VIB Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Genetics, Ghent University, Technologiepark 927, 9052 Gent, Belgium.  
35Institut de Biologie Moléculaire des Plantes du CNRS, Strasbourg 67084, France.  
36BioEnergy Science Center and Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6422, USA.  
37University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.  
38The Ohio State University, Columbus, Ohio 43210, USA.  
39Institut Jean-Pierre Bourgin, UMR1318, Institut National de la Recherche Agronomique, 78026 Versailles cedex, France.  
40Université de Picardie, Amiens 80039, France.  
41Plant Gene Expression Center, University of California Berkeley, Albany, California 94710, USA.  
42Illinois State University and DOE Great Lakes Bioenergy Research Center, Normal, Illinois 61790, USA.  
43Sabanci University, Istanbul 34956, Turkey.  
44Unité de Recherche en Génomique Végétale: URGV (INRA-CNRS-UEVE), Evry 91057, France.  
45USDA-ARS/Donald Danforth Plant Science Center, St Louis, Missouri 63130, USA.  
46Present address: The School of Plant Molecular Systems Biotechnology, Kyung Hee University, Yongin 446-701, Korea.