

SLICING THE wheat genome

By Kellye Eversole,¹ Catherine Feuillet,^{1,2} Klaus F. X. Mayer,³ Jane Rogers¹

This year celebrates the 100th anniversary of the birth of Norman Borlaug, the Nobel Prize-winning plant geneticist who, through his contribution to the “green revolution,” reminds us of the importance of applying scientific knowledge to develop crop varieties. This is even more important today as we face a rapidly expanding global population, climate change, and the need to keep agricultural efforts sustainable while minimizing environmental impacts. Accessing the fundamental information of crop genomes aids in accelerating breeding pipelines and improves our understanding of the molecular basis of agronomically important traits, such as yield and tolerance to abiotic and biotic stresses.

Obtaining a reference sequence of the genome of bread wheat (*Triticum aestivum*), the staple food for 30% of the world’s population, is a scientific challenge. Wheat’s hexaploid genome was formed from multiple hybridization events between three different progenitor species (comprising three individual subgenomes: A, B, and D). This resulted in a large—five times that of humans—and highly redundant genome with more than 80% of the genome consisting of repeated sequences. For these reasons, a reference sequence—a contiguous sequence ordered along the chromosomes—cannot be generated by using whole-genome shotgun sequencing approaches with current high-throughput short read technologies. To overcome this complexity, the International Wheat Genome Sequencing Consortium (IWGSC) developed a strategy of physical mapping and sequencing the individual chromosomes and chromosome arms of the bread wheat genome. In this special issue of *Science*, four Research Articles are presented in full online (www.sciencemag.org/extra/wheatgenome), with abstracts in print on p. 286 and a News story on p. 251. These papers present major advances toward obtaining a reference sequence and enhancing our understanding of the bread wheat genome.

The IWGSC produced a survey of the gene content and composition of all 21 chromosomes and identified 124,201 gene loci, with more than 75,000 positioned along the chromosomes. Comparing the bread wheat gene sequences with gene repertoires from its closest extant relatives (representing the species that donated the A, B, and D progenitor genomes) showed limited gene loss during the evolution of the hexaploid wheat genome but frequent gene duplications after these genomes came together. Gene expression patterns revealed that none of the subgenomes dominated gene expression.

Choulet *et al.* describe the sequencing, assembly, annotation, and analysis of the reference sequence of the largest wheat chromosome, 3B, which at nearly 1 gigabase is more than seven times larger than the entire sequence of the model plant *Arabidopsis thaliana*. Relying on a physical map derived from the chromosome 3B-specific bacterial artificial chromosome (BAC) library (1), more than 8000 BAC clones were sequenced and assembled into a pseudomolecule—a nearly complete representation of the entire chromosome. This high-quality, ordered sequence revealed a partitioning into distinct regions along the chromosome, including distal segments that are preferential targets for recombination, adaptation, and genomic plasticity. Many inter- and intrachromosomal duplications were also observed, illuminating the structural and functional redundancy of the wheat genome. This sequence, which can be anchored to the genetic and phenotypic maps, will aid breeders by increasing the pace and simplifying the process of identifying and cloning genes underlying agronomic traits.

Marcussen *et al.* used the IWGSC chromosome survey sequences to analyze the timing and phylogenetic origin of the diploid genomes that have come together to form the A, B, and D subgenomes of bread wheat. They unravel ancient hybridization events in the wheat lineage and reveal that the ancestral A and B genomes diverged from a common ancestor ~7 million years ago. They also show that the D genome was formed through homoploid hybrid speciation—hybridization that

does not result in a genome duplication event—between relatives of the A and B genomes 1 million to 2 million years later.

Pfeifer *et al.* address inter- and intragenomic gene expression regulation within a polyploid genome by providing an in-depth analysis of the transcriptional landscape of the developing wheat grain. They show that the transcriptional network delineates a complex and highly orchestrated interplay of the individual wheat subgenomes and identify transcriptional active or inactive domains along the chromosomes that might indicate epigenetic control of grain development.

Together, these Research Articles explore multiple dimensions of the 17-gigabase wheat genome and pave the way toward achieving a full reference sequence to underpin wheat research and breeding.

REFERENCE

1. E. Paux *et al.*, *Science* **322**, 101–104 (2008).

10.1126/science.1257983



¹International Wheat Genome Sequencing Consortium, 5207 Wyoming Road, Bethesda, MD 20816, USA. ²Bayer CropScience, Innovation Center, 3500 Paramount Parkway, Morrisville, NC 27560, USA. ³Plant Genome and Systems Biology, Helmholtz Zentrum Munich, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany.

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Lists of authors and affiliations are available in the full article online.

Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de

Read the full article at <http://dx.doi.org/10.1126/science.1251788>

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen, Simen R. Sandve,* Lise Heier, Manuel Spannagl, Matthias Pfeifer, The International Wheat Genome Sequencing Consortium,† Kjetill S. Jakobsen, Brande B. H. Wulff, Burkhard Steuernagel, Klaus F. X. Mayer, Odd-Arne Olsen

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The list of author affiliations is available in the full article online.*Corresponding author. E-mail: simen.sandve@nmbu.no †The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.

Read the full article at <http://dx.doi.org/10.1126/science.1250092>



Triticum monococcum

Triticum carthlicum

Ancestral wheat

Wheat varieties and species (shown) believed to be the closest living relatives of modern bread wheat (*T. aestivum*). Multiple ancestral hybridizations occurred among most of these species, many of which are cultivated, and along with *T. aestivum* represent a dominant source of global nutrition.



Triticum boeoticum

Triticum polonicum L.



Triticum macha

Triticum dicoccoides var. *araraticum*



Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer, Karl G. Kugler, Simen R. Sandve, Bujie Zhan, Heidi Rudi, Torgeir R. Hvidsten, International Wheat Genome Sequencing Consortium, * Klaus F. X. Mayer, Odd-Arne Olsen†

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type-specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type- and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The list of author affiliations is available in the full article online. *The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials. †Corresponding author. E-mail: odd-arne.olsen@nmbu.no

Read the full article at <http://dx.doi.org/10.1126/science.1250091>

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet, * Adriana Alberti, Sébastien Theil, Natasha Glover, Valérie Barbe, Josquin Daron, Lise Pingault, Pierre Sourdille, Arnaud Couloux, Etienne Paux, Philippe Leroy, Sophie Mangenot, Nicolas Guilhot, Jacques Le Gouis, Francois Balfourier, Michael Alaux, Véronique Jamilloux, Julie Poulain, Céline Durand, Arnaud Bellec, Christine Gaspin, Jan Safar, Jaroslav Dolezel, Jane Rogers, Klaas Vandepoele, Jean-Marc Aury, Klaus Mayer, Hélène Berges, Hadi Quesneville, Patrick Wincker, Catherine Feuillet

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaptation. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

The list of author affiliations is available in the full article online. *Corresponding author. E-mail: frederic.choulet@clermont.inra.fr

Read the full article at <http://dx.doi.org/10.1126/science.1249721>

PHOTOS: SUSANNE STAMP; ERNST MERZ/ETH ZÜRICH