

The International Wheat Genome Sequencing Consortium

IWGSC¹, Rudi Appels², Ute Baumann³, Hikmet Budak⁴, Isabelle Caugant¹, Kellye Eversole¹, Etienne Paux⁵ and Sébastien Prud'homme⁶

(1) Lee's Summit, MO, USA; (2) University of Melbourne, Agribio, Australia; (3) University of Adelaide, Australia; (4) Montana State University, MT, USA; (5) National Institute for Agricultural Research (INRA), Clermont-Ferrand, France; (6) BIOGEMMA, Chappes, France

www.wheatgenome.org

Bread wheat, the staple food for 35% of the world's population, is the last major crop species to benefit from a high quality reference genome sequence. The International Wheat Genome Sequencing Consortium (IWGSC), with 2,400 members in 68 countries, is an international, collaborative consortium, established in 2005 by a group of wheat growers, plant scientists, as well as public and private breeders who were committed to delivering high quality genomic tools and resources for wheat improvement.

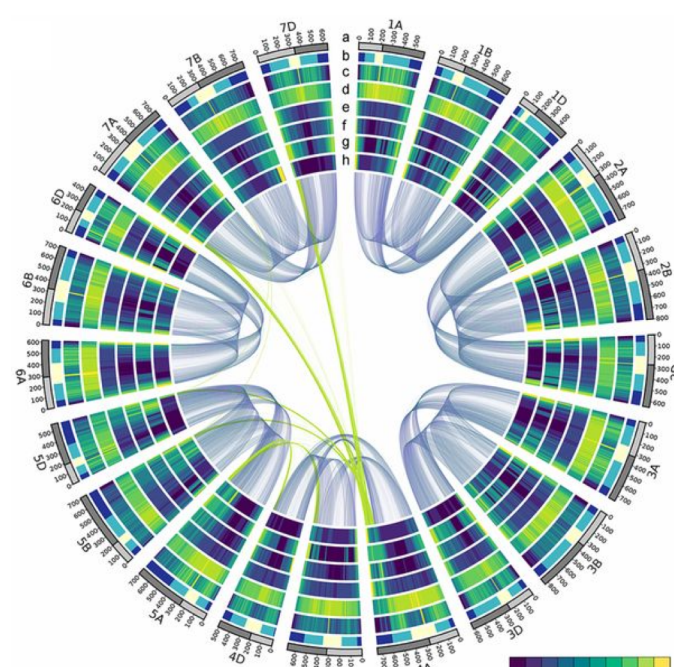
In 2018, the IWGSC completed Phase I with the publication of the first high quality reference sequence of bread wheat, *Triticum aestivum*, cv. Chinese Spring (IWGSC RefSeq v1.0). The reference sequence represents 94% of the hexaploid wheat genome organized in 21 chromosome-like pseudomolecules; identifies the structural features of the wheat genome, including 107,891 high confidence genes in their genomic context, along with more than 4.7 million molecular markers. This work lays the foundation for genomics-based crop improvement in wheat and will underpin future efforts to respond to the challenges imposed by population expansion and a rapidly changing climate.



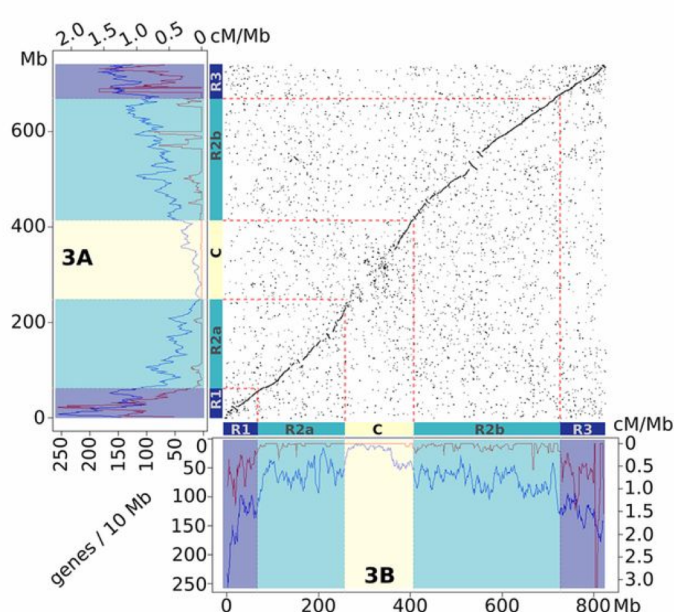
"Shifting the limits in wheat research and breeding using a fully annotated reference genome"
Science. 2018 Aug 17;361(6403)
 doi: 10.1126/science.aar7191

Wheat Genome Structure and Function

A,B,D genome comparisons reveal high levels of sequence collinearity between homeologous chromosomes.



Five chromosomal regions have been defined for all chromosomes. R1 and R3 (purple) corresponding to chromosome extremities are recombination-rich, with high gene- and low TE-densities and contain the highest proportions of duplicated genes and genes expressed in fewest conditions. R2a and R2b (teal blue) regions are intermediate in character between R1/R3 and the gene-poor, recombination-free C regions, where collinearity between homeologs breaks down.

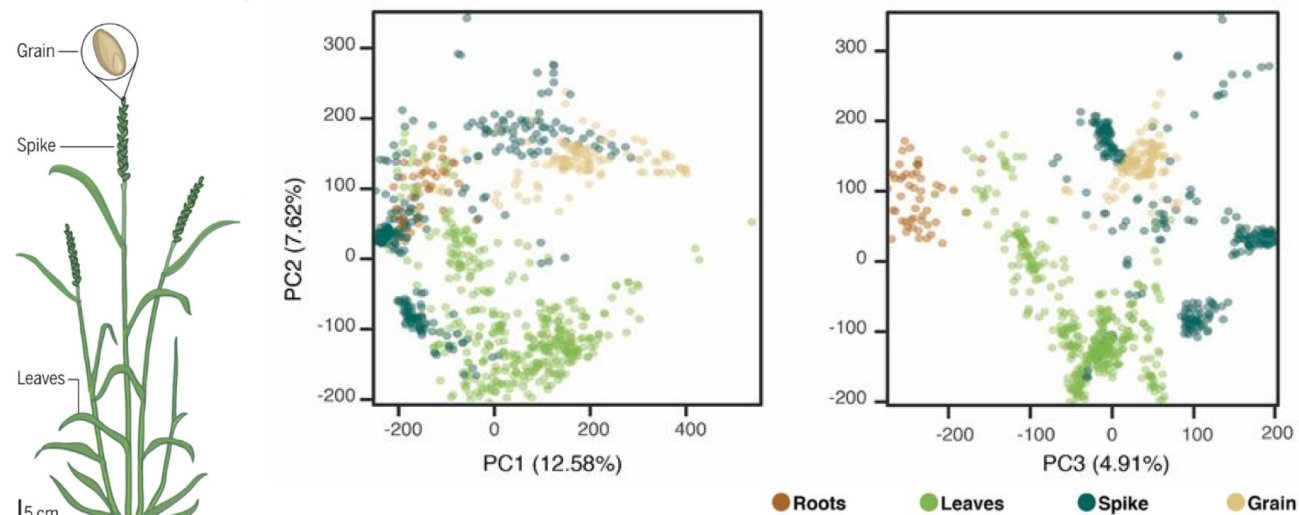


With high recombination, distal R1 and R3 regions provide the genomic environment for creating sequence diversity which gives the basis for the wheat plant's ability to adapt to changing environments.

Legend: Dot-plot visualization of collinearity between homeologous chromosomes 3A and 3B in relation to distribution of gene density and recombination frequency.

Wheat Transcriptome Atlas

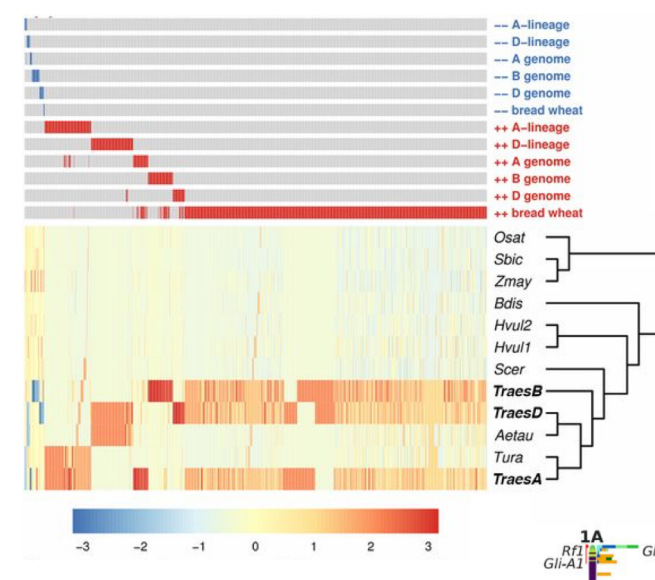
A total of 850 RNA-seq samples derived from 32 tissues at different growth stages and/or challenged by different stress treatments were mapped to IWGSC RefSeq v1.0



Legend: Principal component (PC) analysis plots for similarity of overall transcription. Tissue, rather than growth stage or stress, as the main factor driving differential expression between samples.

Gene Families

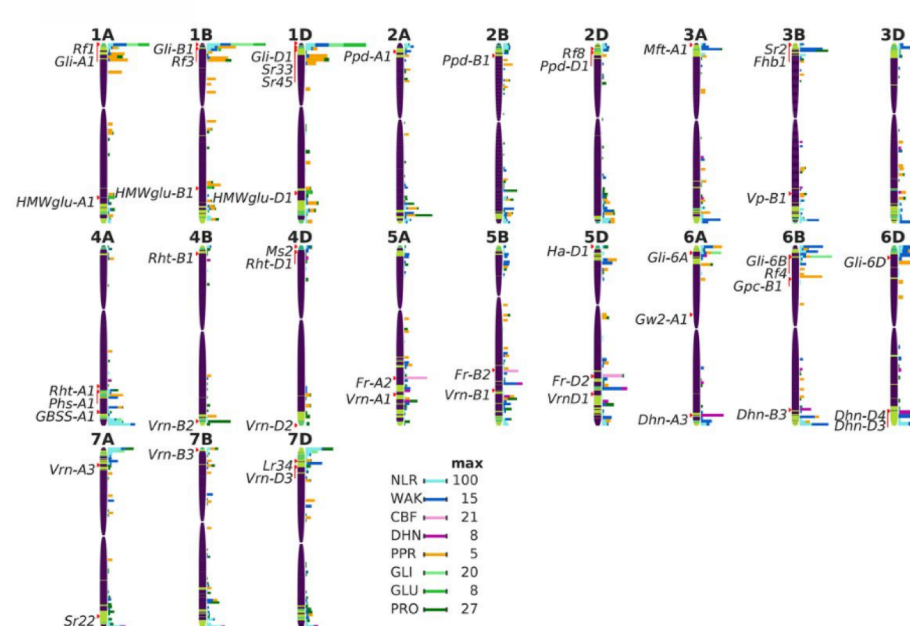
Among the 8592 expanded wheat gene families (33% of all families), 6216 were expanded in all three A, B, and D sub-genomes (24%; either shared with the wild ancestor or specific to bread wheat). Another 1109 were expanded in only one of the wheat sub-genomes, and 2102 gene families were expanded in either the A or the D genome lineages.



Legend: Heatmap of expanded and contracted gene families.

Many gene families with high relevance to wheat breeding and improvement were among the expanded groups, and their genomic distribution was analyzed in greater detail.

Legend: Genomic distribution of gene families associated with adaptation to biotic (light and dark blue) or abiotic stress (light and dark pink), RNA metabolism in organelles and male fertility (orange), or end-use quality (light, medium, and dark green).



Next Step: IWGSC Phase II

With the completion of IWGSC RefSeq v1.0, the Consortium is now in phase II activities:

- IWGSC Exome Array based on the reference sequence (see poster #0953)
- Wheat pan-genome based on de novo sequencing and assembly of multiple wheat genomes (landraces and elite varieties) that cover the breadth of wheat diversity
- Completion of the IWGSC Gold Standard through gap closure, sequence maintenance and updates, as well as integration of manual and functional annotation of the reference sequence
- User-friendly, integrated databases and tools to benefit public breeders and the full range of industry partners, from SME to large companies

All IWGSC data are publicly available at URGI, INRA Versailles, for BLAST searches, download and browser format. <https://wheat-urgi.versailles.inra.fr>

Thanks to IWGSC sponsors

