

International Wheat genome Sequencing Consortium

Science Publication – 18 July 2014

“A chromosome-based draft sequence of the hexaploid bread wheat genome”

Press Kit

- IWGSC press release	2
- Questions & Answers	3
- Comments	6
- Quotes	8
- About IWGFC	10
- Contacts	11

Genetic blueprint of bread wheat genome unveiled

Last step before full genome sequence

Bethesda, Maryland, USA – July 18, 2014

The International Wheat Genome Sequencing Consortium (IWGSC) published today in the international journal *Science* a draft sequence of the bread wheat genome. The chromosome-based draft provides new insight into the structure, organization, and evolution of the large, complex genome of the world's most widely grown cereal crop.

The genetic blueprint is an invaluable resource to plant science researchers and breeders. For the first time, they have at their disposal a set of tools enabling them to rapidly locate specific genes on individual wheat chromosomes throughout the genome. Jorge Dubcovsky, Professor at the University of California Davis, USA, says that these results “have been a fantastic resource for our laboratory. The development of genome specific primers, which used to take several weeks of work, can now be done in hours. Mapping of any sequence to the specific chromosome arm can now be done *in silico* in minutes. In addition to the acceleration of day to day work in wheat genetics, this resource has made possible analyses and discoveries at the genome level that were not possible before.”

The draft sequence is a major landmark towards obtaining a complete reference sequence of the hexaploid bread wheat genome, the ultimate aim of the International Wheat Genome Sequencing Consortium. In the same issue of *Science*, another article presents the first reference sequence for the largest chromosome, 3B. This establishes a proof of concept and a template for sequencing the remaining chromosomes. As of today, researchers in the IWGSC estimate that the full genome sequence will be available within three years.

“With the draft gene sequence for each of the bread wheat chromosome and the first reference sequence of chromosome 3B, we have reached a great milestone in our roadmap,” said Catherine Feuillet, IWGSC co-chair. “We know now the way forward to obtain a reference sequence for the 20 remaining chromosomes and we hopefully will be able to find the resources to achieve this in the next three years.”

With a chromosome-based full sequence in hand, plant breeders will have high quality tools at their



disposal to accelerate breeding programs and to identify how genes control complex traits such as yield, grain quality, disease, pest resistance, or abiotic stress tolerance. They will be able to produce a new generation of wheat varieties with higher yields and improved sustainability to meet the demands of a growing world population in a changing environment.

The draft sequence is already providing new insights into the history and evolution of the wheat genome and genes involved in grain development, as exemplified in two additional publications appearing in the same issue of *Science*.

Wheat is a major dietary component for many populations across the world. Grown on more land than any other crop, more than 215 million hectares of wheat are harvested annually to generate a world production of almost 700 million tons, making it the third most produced cereal after maize and rice. It is the leading source of vegetable protein in human food, having a higher protein content than either maize or rice. The wheat plant is highly versatile due to its ability to grow in a wide range of environments. Wheat grain is easily stored and can be converted readily into flour for making numerous varieties of high quality edible food.

About the IWGSC:

The IWGSC, with more than 1,000 members in 57 countries, is an international, collaborative consortium, established in 2005 by a group of wheat growers, plant scientists, and public and private breeders. The goal of the IWGSC is to make a high quality genome sequence of bread wheat publicly available, in order to lay a foundation for basic research that will enable breeders to develop improved varieties.

More information can be found at

www.wheatgenome.org

Contact:

Isabelle Caugant

communications@wheatgenome.org

+32 (0)484 750 634



Who participated?

The IWGSC article has almost 100 authors, representing 23 laboratories from 15 different countries. They are made up of groups who participated in generating the data through providing wheat lines, flow sorting DNA and sequencing and bioinformaticians and wheat geneticists that analyzed the data.

What did they do?

They produced the first survey of the gene content and composition - also called “draft sequence” - of the 21 individual chromosomes of bread wheat. Out of 124,201 sequences that could be recognized as genes or gene loci, more than 75,000 were positioned along the chromosomes.

They also produced and/or used whole genome shotgun sequences to identify genes in species of diploid and tetraploid wheat that are related to the ancestors of hexaploid bread wheat. Finally, they measured gene expression in five different tissues of wheat plants taken at three stages of development and studied how genes from different chromosomes behave in a hexaploid context.

The data that have been generated provide a unique resource for accelerating gene mapping and marker development in wheat breeding, as well as for analyzing how wheat genes are used and how they have evolved.

How did they do it?

Using DNA isolated from individual wheat chromosomes, they produced sequence information covering each base on average at least 30 times by using high throughput short read DNA sequencing technologies (so called ‘Next Generation Sequencing’). The sequences were then assembled into short fragments of unique DNA and were assigned positions along the chromosomes by applying an approach called the GenomeZipper to produce a pseudo-order by aligning the contigs against wheat genetic maps and other sequenced grass genomes.

Are the results publicly available?

Yes.

A central IWGSC repository has been established within the national plant bioinformatics database in France (the Unité de Recherche Génomique (URGI) at Institut National de la Recherche Agronomique (INRA)) to provide public access to wheat sequence data and other IWGSC resources, for example, physical maps and marker data. All data from the survey sequence initiative are available for the scientific community at this address:

<https://urgi.versailles.inra.fr/download/iwgsc/> or
<http://wheat-urgi.versailles.inra.fr/Seq-Repository/>

The raw sequence reads for each chromosome have also been deposited in the Short Read Archive (SRA) at the European Bioinformatics Institute and the assembled sequence data have been integrated into Ensembl Plants at EBI at the following address:

http://plants.ensembl.org/Triticum_aestivum/

How are these results significant and how will they be used?

For scientists: the draft sequence provides a unique resource for rapidly localizing specific genes on the chromosomes and for comparative genomics analysis with related species to study the evolution of the gene content in the wheat lineage. It also provides a foundation for analyzing the expression of homoeologous genes in a large polyploid genome.

For plant breeders: The draft sequence provides the putative location of about 75,000 genes along the wheat chromosomes and a substrate to design genome-specific markers through sequence capture and resequencing activities. Breeders can then use marker sets to map genes and to deploy marker assisted breeding and genomic selection schemes.

How much did that cost?

IWGSC estimates the project cost in the order of €200,000 for DNA preparation and the sequencing which was carried out by seven laboratories around the world. Salaries and any costs associated with bioinformatics and analyses are not included in this estimate.

What's the next step?

The IWGSC's focus is to continue its efforts to obtain a complete wheat genome sequence, also called a reference sequence. The wheat genome contains over 80% of repetitive sequences that cannot be assembled unambiguously from the short stretches of sequence of 100 to 150 bases generated with the current high throughput sequencing technologies. As a result, the survey sequences are still very fragmented. They do not provide long stretches of sequence that include regulatory sequences and several genes in a single sequence. Moreover, they provide only partial information about the order and orientation of the contiguous sequences and thus the genes and their surrounding sequences that make up the chromosomes. The information available for positional cloning of genes underlying agronomically important traits is therefore incomplete.

In the same issue of *Science*, another article reports the completion of the first reference sequence for the largest chromosome, 3B. This establishes a proof of concept and a template for sequencing the remaining 20 chromosomes.

When will the wheat genome sequence finally be completed?

With the demonstration that it is necessary and possible to sequence the wheat genome to the high quality as has been achieved for wheat chromosome 3B, the IWGSC has developed a three-year plan to complete the reference sequences for the remaining 20 chromosomes. Sequencing of some chromosomes is already underway and the whole genome could be completed by 2017 if the remaining funding requirements were to become available immediately.

The aim is to obtain a high quality sequence that provides an accurate representation of the structure and organization of sequences along individual chromosomes, enabling the identification of the positions of genes, regulatory elements, repetitive elements, sequence-based markers, and other features.

Why does it take so long?

Obtaining a high quality reference sequence of the genome of bread wheat is a scientific challenge. With a size of 17 Gb, the wheat genome is more than five times larger than the human genome. In addition, the bread wheat genome comprises 21 chromosomes originating from three individual subgenomes that contain highly similar gene sets and have a repeat content of over 80%. Because of these issues, the IWGSC decided that the only approach that would deliver a high quality reference was to reduce the complexity and follow a strategy similar to that used for other high quality reference genomes - such as human, mouse, zebrafish, *Arabidopsis* and rice - namely, sequencing each chromosome. It has been a challenge to secure funding in part because investment into wheat research is generally lower than in other major crops such as maize, despite the importance of wheat as a major source of human food.

Why is it useful to sequence the wheat genome?

The world is facing enormous challenges with a human population projected to rise to over 9 billion by 2050. Food production will need to increase by over 50% without expanding land use in the face of a changing climate and with dwindling availability of fertilisers, water and effective pest treatment. To produce sufficient wheat for the human population in the future, there is an urgent need to develop new wheat varieties with higher yield, better resistance to diseases and pests, and tolerance to abiotic stresses such as drought, high salinity or high aluminium content of the soils.

Once the reference genome sequence is completed, breeders will have at their disposal tools to identify genes and regulatory elements underlying complex traits and accelerate improvement through genomics assisted breeding and biotechnology. In particular, they will be able to understand the interplay between sets of similar genes that reside in the genome and can be regulated in different ways. Using this information, breeders will be able to produce a new generation of wheat varieties with higher yields and improved sustainability to meet the demands of a growing world population in a changing environment. The draft genome sequence provides a useful first step towards this goal.

Questions & Answers

Wasn't the wheat genome sequenced two years ago?

In 2012, a group in the United Kingdom reported the production of sequence information that represented approximately 5-fold coverage of the wheat genome, covering up to 70% of the non-repetitive regions, and allowed them to identify approximately 95,000 gene sequences. They made putative assignments of genes to each of the A, B, and D subgenomes, based on similarity to two diploid wheat species that are relatives of wheat A and D genomes and that had been sequenced by two Chinese groups. They were not able to assign genes to individual chromosomes, however and they did not provide information that allows breeders to identify differences between genes that lie on chromosomes within each of the A, B and D subgenomes.



Bread wheat originates from three ancestral diploid species related to *T. urartu*, *Ae. speltoides* and *Ae. tauschii* (upper 3 pictures) that consecutively hybridized into tetraploid durum (pasta) wheat (lower left quadrant) and hexaploid bread wheat (lower right quadrant).

In what way have the data from the draft sequence helped you in your work?

Julien Bonneau - Postdoctoral Researcher, University of Melbourne, Australia

I have been a user of the IWGSC databases and would recommend its use to any scientist working on wheat, specifically when one wants to increase marker density and gain access to gene-based markers in positional cloning and gene discovery projects. These databases, along with those related to the 3B physical map and reference sequence, have allowed us to get into candidate gene discovery to improve wheat grain yield where plants are grown in challenging environmental conditions such as drought and/or heat. Without this information we would be far behind and struggling to make progress.

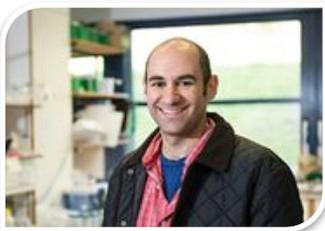
In addition, our group has been able to identify a large gene family (up to 20 genes) potentially involved in Iron and zinc uptake. This information is extremely precious not only in the context of increasing production but even more with regards to grain quality and to combat malnutrition.

In fact without these resources we wouldn't have been able to discover these genes as fast as we did! It's really important to make fast progress in understanding mechanisms in micronutrient uptake to produce crops that would benefit human health in the future.



Jorge Dubcovsky - Professor, University of California Davis, and HHMI-GBMF researcher, United States

The IWGSC chromosome shotgun sequences and associated data have been a fantastic resource for our laboratory. Development of genome specific primers, which used to take several weeks of work and painful sequence of the three genomes can now be done in hours. Mapping of any sequence to the specific chromosome arm can now be done in silico in minutes. In addition to the acceleration of day to day work in wheat genetics, this resource has made possible analyses at the genome level that were not possible before. It provides a reference for RNASeq studies, gene capture, promoter identification, evolution of homoeologs, etc.



Cristobal Uauy - Project Leader, John Innes Centre, Norwich, United Kingdom

The IWGSC CSS data is making a massive difference in how we do wheat genetics today. It has opened up new areas of science in wheat, for example, by providing access to promoter sequences across multiple genomes. It has enabled us to establish high-throughput methods for marker development allowing us to quickly generate genome-specific assays for fine mapping and breeding purposes. The CSS data is also proving fundamental in our efforts to generate an *in silico* TILLING mutant collection in bread and pasta which will further accelerate functional genomics in wheat.



Jan Gielen - Senior Scientist Applied Genomics, Syngenta, France

Marker assisted breeding for wheat has shifted gears since the resources developed through the IWGSC initiative became available. The chromosome survey sequences revealed instrumental for increasing the throughput of various marker technologies and marker applications at ever lower cost in a hexaploid crop like wheat.

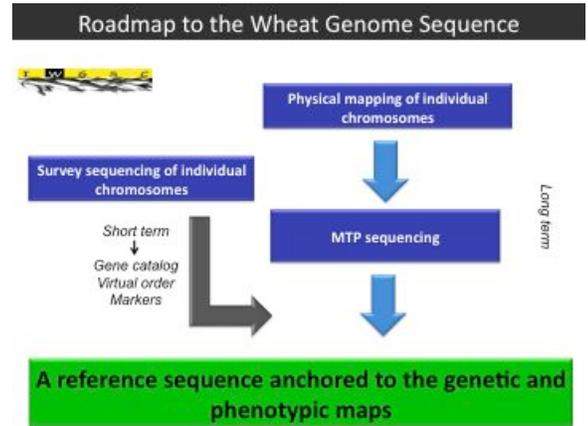
How does the publication of the draft sequence impact the IWGSC roadmap?



Kellye Eversole - IWGSC director

There are three key milestones in the IWGSC roadmap: (1) achieving survey sequences of the 21 bread wheat chromosomes; (2) developing physical maps that can serve as a substrate for sequencing; and (3) completing the reference sequence for each of these chromosomes.

This publication describes the achievement of the first milestone.



Which methodology was used to produce the draft sequence?



Jane Rogers - IWGSC co-director

The draft sequence was generated from DNA of the bread wheat cultivar Chinese Spring that was isolated from individual wheat chromosome arms, or the whole chromosome in the case of chromosome 3B. The chromosomes were separated using flow-sorting by Jaroslav Dolozel's laboratory at the Institute of Experimental Botany, Olomouc, in the Czech Republic.

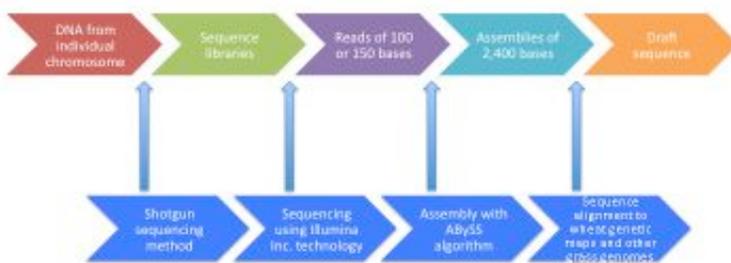
The shotgun sequencing method was used to fragment the chromosomal DNA into pieces with an average size of 500 base pairs and to create libraries for sequencing using the paired end sequencing technology of Illumina, Inc. Sequence reads with a length of 100 or 150 bases were produced that covered the chromosomes to a depth of more than 30-fold.

The sequence reads were trimmed as necessary to remove poor quality bases. Then, bioinformaticians at The Genome Analysis Centre in Norwich, United Kingdom, assembled them using the assembly algorithm ABySS that was developed at the Michael Smith Genome Sciences Centre, in Vancouver, Canada.

The sequences were assembled into lengths (contigs) with an average size of 2,400 bases that, in total, covered 10.2 Gb, or approximately 60% of the wheat genome. The assemblies contained very little information to position them along the chromosomes, so Klaus Mayer and colleagues at the Helmholtz Institute, Munich, in Germany, used an approach called the GenomeZipper to produce a pseudo-order by aligning the contigs against wheat genetic maps and other sequenced grass genomes.

The sequences represented in the assemblies contain a large proportion of the wheat genes, whilst the most repeated sequences that make up large tracts of DNA between the genes, are not captured in the draft sequence.

Methodology used to produce the draft sequence



From the authors



Catherine Feuillet - IWGSC co-chair

“As a co-chair of the IWGSC, I am really delighted to witness the success of the international team effort that led to the survey sequence and the achievement of one of our key milestone. We launched this effort, at the end of 2010 and assembled the most competent laboratories in wheat genomics to tackle the task of producing the survey sequence. Such a collective effort around a focused objective defined by the IWGSC coordinating committee is a typical illustration of the IWGSC philosophy and it is rewarding to see how successful this is.”



Jane Rogers - IWGSC co-director

“The draft sequence represents a significant milestone in the sequencing of the wheat genome, providing an important set of tools to differentiate between genes residing on the three sub-genomes and large numbers of new markers that can be used to pursue genes underlying important traits. Publication of the draft genome sequence alongside the reference sequence of chromosome 3B, however, provides a timely reminder that structural complexity not revealed by the draft plays a key role in genome regulation and that wheat scientists and breeders will require a reference sequence of the whole genome to fully access and apply wheat genomics to wheat crop improvement.”



Jaroslav Doležel - Institute of Experimental Botany, Centre of Plant Structural and Functional Genomics, Czech Republic

“Publication of the wheat genome blueprint and the impact this advance makes are a terrific reward for many years of hard and systematic work of our team. It is great to see that the chromosome-based strategy we have developed overcomes intricacies of the complex wheat genome to deliver the knowledge and molecular tools for wheat improvement. This is the best motivation for us to continue together with other members of the Consortium.”



Klaus Mayer - German Research Center for Environmental Health, Helmholtz Zentrum München, Germany

“The new insights into the biology of the wheat genome now enable to identify genes and markers for breeding more rapidly. These are fundamental requirements for facing the challenges imposed by increased food demands, plant diseases and a changing climate.”



Curtis Pozniak - Crop Development Centre, University of Saskatchewan, Canada

“The release of the chromosomal draft of the wheat genome sequence will accelerate gene discovery in wheat, and pave the way for development of high-throughput tools to improve breeding of complex traits such as yield, insect and disease resistance, and end-use quality.”

Quotes



Kelly Eversole - IWGSC executive director

“As the executive director, it is really exciting to see the hard work of so many IWGSC members rewarded first with achieving a major milestone and second with a publication in the prestigious journal, *Science*. We began the consortium in 2005 with a handful of scientists and support from a few growers and industry partners. As we grew to become the international organization we are today, we developed our roadmap with guidance from industry and wheat scientists and, thus, with a view towards delivering a sequence that will be useful for end users. Along the way, many thought it would be impossible to reach any of our objectives and encouraged us to lower our standards. We chose instead to maintain a constant focus on application, the pursuit of high quality tools and resources for breeders and wheat scientists, and on building a foundation for sequencing that would permit the integration of advancements in sequencing technologies. The milestones described in *Science*, specifically the completion of the IWGSC Chromosome Shotgun Sequences and the first high quality reference sequence of a wheat chromosome (chromosome 3B), are a testament to the resounding success of our process and strategy as well as a reminder that it is still possible to achieve great things as a group.”

From others



H el ene Lucas - International Scientific Coordinator of the Wheat Initiative

“The achievement by the IWGSC of the survey sequence of the wheat genome and of the reference sequence of the 3B chromosome is a very important milestone for the completion of the reference sequence of the whole genome, which the Wheat Initiative has identified as a top priority for wheat research and improvement.”

Andrew Sharpe, Research Officer at the National Research Council Canada

“The achievement of the chromosome-based draft is a technical *tour de force* and highlights the improvements in sequencing and bioinformatics over the last few years. The work has provided a fascinating glimpse into the evolution of the wheat genome. However, the full story will only be established with the development of a reference genome sequence.”

About the IWGSC

It all started in 2005 when a group of growers, breeders, and scientists launched the International Wheat Genome Sequencing Consortium to advance wheat improvement and ensure profitability throughout the wheat value chain. At that time, genomic resources for sequencing the wheat genome and the pace of variety improvement were lagging significantly behind other major crops such as maize and rice.

To change this paradigm, the IWGSC is dedicated to producing a reference sequence of the bread wheat genome for accelerating molecular breeding, better understanding of the molecular basis of key agronomic traits, and knowledge of the structure and function of the wheat genome.

As of 2014, the IWGSC has projects in 23 countries and more than 1,000 members in 57 countries.

Resources available for all

Since its inception, the ultimate goal of the IWGSC has been to produce publicly-available resources that are relevant for improving wheat species used for food and feed consumption.

The aim is to generate a resource that provides an accurate representation of the structure and organization of the sequence along individual chromosomes, enabling rapid localization of genes, regulatory elements, repetitive elements, molecular markers, and the linkage of these features with phenotypes.

A milestone-based strategy

A Coordinating Committee comprised of 63 members representing universities, public research and grower organizations, as well as small and large wheat breeding and seed companies define the IWGSC strategy and roadmap.

To ensure the rapid delivery of tools to wheat breeders and scientists, the IWGSC establishes strategic plans with short- and mid-term goals, defines areas of coordination, facilitates and coordinates research projects and funding efforts at the national and international levels, develops and supports the design of research proposals, provides a framework for the establishment of common guidelines, protocols, and resources, and organizes scientific meetings and workshops.

Using its milestone-based, adaptable strategy, the IWGSC regularly provides new tools and resources to breeders while working towards the ultimate goal of obtaining a high quality reference genome sequence for bread wheat.

This overall strategy ensures the immediate availability of significant outputs for wheat breeders and the wheat industry at large in parallel to continued advancements in basic research on the wheat genome.

Leadership Team

Kellye Eversole, executive director, and Jane Rogers, deputy executive director, manage the day-to-day scientific and strategic activities of the IWGSC. They are supported by six co-chairs: Rudi Appels (Murdoch University), Jan Dvorak (University of California, Davis), Catherine Feuillet (Bayer CropScience), Bikram Gill (Kansas State University), Beat Keller (University of Zurich), and Yasunari Ogihara (Yokohama City University).

General membership in the consortium is open to anyone and general business meetings are open to the public.

The IWGSC organizes at least three scientific conferences and workshops each year that usually are held in conjunction with other relevant meetings such as the annual Plant and Animal Genome Conference and the International Triticeae Mapping Initiative meeting.



Contacts

IWGSC media contact

Isabelle Caugant
communications@wheatgenome.org
Tel: +32 484 750 634

Media contact at IWGSC members

In Germany: Susanne Eichacker
 Helmholtz Zentrum München
 German Research Center for Environmental Health
 susanne.eichacker@helmholtz-muenchen.de
 Tel: +49 89 3187 3117

In the Czech republic Jan Martinek
 Head of Media Communications Division
 The Czech Academy of Sciences
 martinek@kav.cas.cz
 Tel.: +420 221 403 423

In Canada Jennifer Thoma
 Media Relations Specialist
 University of Saskatchewan
 Jennifer.thoma@usask.ca
 Tel.: +1 306 966 1851

International
Wheat Genome
Sequencing
Consortium

www.wheatgenome.org