

From genome to gene: a new epoch for wheat research?

Meng Wang, Shubin Wang, and Guangmin Xia

The Key Laboratory of Plant Cell Engineering and Germplasm Innovation, Ministry of Education, School of Life Sciences, Shandong University, Jinan 250100, P.R. China

Genetic research for bread wheat (*Triticum aestivum*), a staple crop around the world, has been impeded by its complex large hexaploid genome that contains a high proportion of repetitive DNA. Recent advances in sequencing technology have now overcome these challenges and led to genome drafts for bread wheat and its progenitors as well as high-resolution transcriptomes. However, the exploitation of these data for identifying agronomically important genes in wheat is lagging behind. We review recent wheat genome sequencing achievements and focus on four aspects of strategies and future hotspots for wheat improvement: positional cloning, 'omics approaches, combining forward and reverse genetics, and epigenetics.

Progress towards the acquisition of the wheat genome sequence

Since the release of the *Arabidopsis* (*Arabidopsis thaliana*) genome sequence in 2000 an ever-growing number of plant genomes have been sequenced, and these have revolutionized our understanding of plant biology and facilitated the improvement of crops [1–3]. By contrast, the acquisition of the genome sequence of bread wheat (*Triticum aestivum*) has been hampered by three major features distinguishing it from the genomes of other plants such as *Arabidopsis*: (i) the ~17 Gb giant size, which is approximately sixfold larger than the human genome and 125-fold larger than that of *Arabidopsis*, means that initial attempts of genome sequencing were costly; (ii) the extensive stretches of repetitive DNA (>80%) make sequence assembly challenging, not only because they contain multiple highly homologous sequences but also because the transposable element (TE) bursts create numerous non-colinear genes relative to grass models, hindering their anchoring to the 'backbone'; (iii) bread wheat ($2n = 6x = 42$) is a hexaploid species with an AABBDD genome, derived from two amphiploidization events: the first hybridization forming the tetraploid wheat species ($2n = 4x = 28$, genome AABB) was between the *Triticum urartu* ($2n = 2x = 14$, genome AA) and presumably *Aegilops speltoides*, belonging to the section Sitopsis ($2n = 2x = 14$, genome SS); the second hybridization was between the tetraploid wheat and

Aegilops tauschii ($2n = 2x = 14$, genome DD) [4,5]. This hexaploid nature of bread wheat leads to problems in differentiating and assigning the highly conserved homeologous genes originating from the A, B, and D subgenomes. Further complicating sequence assembly is the fact that during the course of evolution several inter-chromosome translocations (such as between the 4AL and 5AL, 7BS) have occurred [6]. With the emergence of next-generation sequencing (NGS) technologies, whole-genome shotgun sequencing became faster and cheaper, and provided a feasible approach for the bread wheat genome. Based on this approach, low-coverage ($5\times$), relatively long-read (454) shotgun sequences of model hexaploid cultivar Chinese Spring were produced using the crude sequences of diploid progenitor genomes as a guide for the assembly of thousands of small contigs [7]. This led to the first database covering the hexaploid wheat genome, and this formed a framework for further sequencing of bread wheat, accelerated marker development, and gave a rough estimation of gene content with ~96 000 genes. Nevertheless, reliable assembly of these sequences in such a complicated genome has proved to be impossible [8].

A more systematic approach is focused on acquiring the genome sequences of three diploid ancestors of wheat (*T. urartu*, *Ae. speltoides* and *Ae. tauschii*), which separates the hexaploid genome into A, B, and D genomes. Currently, draft genome sequences of the A and D genome progenitors (*T. urartu* and *Ae. tauschii*) have been acquired using shotgun sequencing, and these provide a new basis for a comparative analysis with the bread wheat genome and insight into evolutionary aspects [9,10]. However, both the progenitor genomes as well as the shotgun sequencing of hexaploid wheat (described above) are fragmentary and incomplete because the raw data were acquired from short reads. Subsequently, a 4 Gb physical map of *Ae. tauschii* using the SNaPshot BAC fingerprinting technology has been generated [11] and the construction of its reference sequence is on the agenda [12].

Based on the earlier construction of cytogenetic stocks of Chinese Spring, such as telosomic and ditelosomic lines that carry chromosome arms as telocentric chromosomes (telosomes) [13], and the technological advances of high-throughput chromosome isolation using flow cytometry by the group of J. Doležel in Olomouc (Czech Republic) [14], a chromosome-by-chromosome strategy has proved to be the optimal approach to obtain the bread wheat genome. This strategy can reduce the sample size and complexity by sequencing a single chromosome/chromosome arm, avoid

Corresponding author: Xia, G. (xiagm@sdu.edu.cn).

Keywords: wheat; whole-genome sequence; agronomically important genes; strategies for gene identification.

1360-1385/

© 2015 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tplants.2015.03.010>

the confounding effects of homeologous sequences, and allows international cooperation to sequence the individual chromosome parts simultaneously in multiple labs. Consequently, the wheat research community formed the International Wheat Genome Sequencing Consortium (IWGSC) aiming to construct the physical map and to obtain a high-quality reference sequence in bread wheat via the chromosome-based approach [15]. The first chromosome to be successfully sorted was the largest of the chromosome complement, namely chromosome 3B. The DNA from a purified preparation of this chromosome was successfully used to construct the BAC library [16] and to produce the physical map [17]. Recently, another paper released the pseudomolecule of chromosome 3B by sequencing the minimal tiling path of the chromosome physical map and ordering 1358 scaffolds along the chromosome (93% of the complete sequence), which shed light on the partitioning pattern of bread wheat chromosome 3B [18]. In addition to the high-resolution sequence of 3B, a chromosome-based draft of the bread wheat genome was presented by IWGSC in which a summary and preliminary analysis of the chromosome-based sequencing were given [19]. Until now the survey sequences of all chromosomes,

the physical maps of 16 chromosomes, and the reference sequence of 3B are available; at the same time the assembly of physical maps of the other five chromosomes and the acquisition of the complete reference sequence of bread wheat are ongoing (see <http://www.wheatgenome.org/> and Table 1).

Genotyping and phenotyping technologies

Along with the progress in genome sequencing, another milestone in wheat genomics is the emergence of high-throughput SNP-based genotyping technologies. Compared with the traditional markers, SNP can be more abundant and polymorphic, which ensures that sufficient numbers of polymorphic loci can be defined among most genomes. Therefore SNP discovery has been one major task in wheat genetics, but has been progressing slowly for a long time. Early approaches for SNP discovery in wheat were achieved by analyzing sequence variations of cDNAs/expressed sequence tags (ESTs) among wheat lines with PCR and Sanger sequencing (see <http://wheat.pw.usda.gov/SNP/new/index.shtml>) but were costly and complicated. Recent application of NGS significantly improved the efficiency and throughput of SNP discovery. As a result,

Table 1. Online resources for wheat research

Web	Description	URLs
The International Maize and Wheat Improvement Center (CIMMYT)	Ordering different accessions of wheat germplasm and stocks	http://www.cimmyt.org/en/
The USDA NSGC		http://www.ars.usda.gov/main/docs.htm?docid=2884
The IPK genebank		http://www.ipk-gatersleben.de/en/dept-genebank/
The Australian Winter Cereal Collection		http://www.dpi.nsw.gov.au/research/centres/tamworth/specialist-research
KOMUGI wheat genetic resources database		http://www.shigen.nig.ac.jp/wheat/komugi/
Chinese crop germplasm resources		http://www.cgri.net
The N.I. Vavilov Research Institute of Plant Industry	Genotyping and phenotyping different accessions of wheat germplasm and stocks	http://www.nationalgenebank.org/en/
TriticeaeCAP		http://www.vir.nw.ru/
The IWGSC	Genome sequence and SNP assay	http://www.triticeaecap.org/
Wheat genomics		http://www.wheatgenome.org/
JCVI wheat genome database	Genome sequence Gene annotation Wheat–rice comparative maps	http://www.cerealsdb.uk.net/cerealgenomics/Index_Home.html
TriFLDB	Gene annotation EST blast Protein domains	http://www.jcvi.org/wheat/index.php
PlaNet		http://trifldb.psc.riken.jp/v3/index.pl
DfCI gene index		http://aranet.mpimp-golm.mpg.de/
GrainGenes	Gene annotation EST blast QTL markers Transcriptome	http://compbio.dfci.harvard.edu/tgi/plant.html
URGI	Gene annotation EST blast Physical and genetic maps Markers and QTL	http://wheat.pw.usda.gov/GG2/index.shtml
CMap	Genetic map	http://wheat-urgi.versailles.inra.fr/
Wheat iSelect SNP microarray	Wheat SNP array	http://ccg.murdoch.edu.au/cmmap/ccg-live/
wDBTF	Wheat transcription factor database	http://129.130.90.211/snp/
PlantCare	A database of plant <i>cis</i> -acting regulatory elements	http://www.appli.nantes.inra.fr:8180/wDBTF/
		http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

Download English Version:

<https://daneshyari.com/en/article/2825930>

Download Persian Version:

<https://daneshyari.com/article/2825930>

[Daneshyari.com](https://daneshyari.com)