THE CROP JOURNAL XX (2017) XXX–XXX

Available online at www.sciencedirect.com

ScienceDirect

Wheat functional genomics in the era of next generation sequencing: An update☆

Meiling Jia[1], *Jiantao Guan*[1], *Zhiwen Zhai*, *Shuaifeng Geng*, *Xueyong Zhang*, *Long Mao*[*], *Aili Li*[*]

National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

ARTICLE INFO

Article history:
Received 20 July 2017
Received in revised form 13 September 2017
Accepted 13 September 2017
Available online xxxx

Keywords:
GWAS
NGS
TILLING
Triticum aestivum

ABSTRACT

Bread wheat is not only an important cereal crop but also a model for study of an allopolyploid plant with a large, highly repetitive genome. Advances in next-generation sequencing (NGS) technology provide needed throughput to conquer the enormous size of the wheat genome. Multiple high quality reference genome sequences will soon be available. Full-scale wheat functional genomics studies are dawning. In this review we highlight the available tools and methodologies for wheat functional genomics research developed with the assistance of NGS technology and recent progress, particularly the concerted effort in generating multiple reference genomes, strategies to attain genome-wide genetic variation, genome-wide association studies, mutant population generation, and NGS-supported gene cloning and functional characterization. These resources and platforms lay a solid foundation for wheat research, leading to a new era of wheat functional genomics that will bridge the gap between genotype and phenotype. Dissection of wheat genomes and gene functions should assist in genomics-assisted selection and facilitate breeding of elite varieties for sustainable agriculture in China and the world.

Contents

☆ Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

* *Corresponding authors.*
E-mail addresses: maolong@caas.cn (L. Mao), liaili@caas.cn (A. Li).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cj.2017.09.003

Please cite this article as: M. Jia, et al., Wheat functional genomics in the era of next generation sequencing: An update, The Crop Journal (2017), https://doi.org/10.1016/j.cj.2017.09.003

1. Introduction

Bread wheat or common wheat (*Triticum aestivum* L.) is one of the most important global foods. Common wheat originated by hybridization between cultivated tetraploid emmer (*T. dicoccum*, AABB) and diploid goat grass (*Aegilops tauschii*, DD) approximately 10,000 years ago [1]. Wheat is one of the foundation species whose cultivation and domestication has been closely associated with the prosperity of agriculture and settled societies. Common wheat became one of the most widely grown crops due to its high yields and nutritional and processing qualities [2]. By 2050 the world population will reach nine billion. The food needs of those people will require significantly increased wheat yield delivered by genetic improvement. This task is greatly assisted by advances in next-generation sequencing (NGS) technologies that will rapidly change wheat functional genomics studies and also make possible faster forward genetics studies. The wheat community has recently generated genome sequences not only for the wheat diploid donors of the A (*T. urartu*) [3] and D (*Ae. tauschii*) [4] genomes, but also the hexaploid wheat genome (*T. aestivum*) [5,6]. More recently, updated genomic versions for *Ae. tauschii* [7] and Chinese Spring became available [8]. Additional experimental approaches, resources, and computational tools are becoming available for gene identification that can be utilized in wheat breeding.

In this review we summarize recent progress on wheat genomics and functional genomics studies that was achieved with the assistance of the NGS technology that focused on whole genome sequencing of hexaploid wheat and its donor species, genomic polymorphism, cloning of genes of agronomic importance, and development of technical platforms. We envisage that the future of wheat functional genomics will be accelerated under the combined applications of new strategies for genetic mapping and new resources for discovering genetic variation. Identification of gene function will contribute significantly to wheat improvement.

2. Sequencing the large genomes of common wheat and its progenitors

A major advantage of NGS technologies is to make draft sequences of genomes, especially the larger ones, more affordable relative to that of traditional technologies. NGS technologies also provide more possibilities to investigate gene structure and expression. Heritable genome variation that underlies important agronomic traits can be identified in a quicker and systematic manner. Advances in NGS technology provide necessary tools to dissect large and complicated genomes such as hexaploid wheat. Even so, the large genome and polyploid nature of common wheat has been a huge challenge, especially the high percentage (80%–90%) of repetitive sequences [9,10]. It is nearly impossible to distinguish sequences from the highly similar homoeologous subgenomes that have diverged over 2.5–6.0 million years [11,12]. Therefore, sequencing the genomes of the diploid progenitors became an alternative and complementary choice, resulting in genome sequences of *Ae. tauschii* and *T. urartu*, the D and A genome donors, respectively. These sequences provide the necessary genetic information to identify homoeologs in common wheat [3,4].

Due to the complexity of hexaploid wheat there were several early efforts to sequence the genome. An initial *de novo* assembly of sequence data was attempted in chromosome 7DS derived from flow-sorted chromosome arms, demonstrating that it was possible to assemble all known 7DS genes [13]. The same approach confirmed the translocation between chromosome arms 7BS and 4AL in the Chinese Spring genome [14], that exemplified the types of genomic changes that occurred during early evolution and domestication [15]. Such an approach was then applied to all wheat chromosome arms with the exception of 3B [6], which was isolated as an intact chromosome, resulting in the first draft genome assemblies for wheat chromosomes. The genomes of two additional cultivars, Opata M85 and W7984, were shotgun-sequenced by Illumina Solexa, albeit with only limited annotation performed on the assemblies [16]. Among these efforts, the chromosome 3B sequence represented the highest quality sequence, which was based on a physical map [17]. A better picture of wheat chromosome structure and functional partitioning was obtained by studying these sequences. Study of genomic diversity among wild and domesticated accessions can reveal genomic regions bearing the signature of selection under domestication [18]. In 2017, a 10.1-gigabase assembly of the 14 chromosomes of wild emmer (*T. turgidum* ssp. *dicoccoides*) was produced using the NRGene genome assembly algorithm and provided a detailed analysis on the gene content, genome architecture, and genetic diversity of this related AB genome donor to common wheat [19].

The complete sequencing of these wheat genomes with the assistance of NGS technology are milestones for wheat biology and provide long needed resources for wheat functional genomics. Despite this, in order to link phenotypic traits to functional genes, wheat researchers had to work on additional technical platforms and resources such as mutant libraries, full-length cDNA clones, and SNP microarrays.

3. NGS-based genotyping reveals the genetic diversity of wheat

Sequencing mRNA pools is one of the most efficient single nucleotide polymorphism (SNP)-discovery approaches

Please cite this article as: M. Jia, et al., Wheat functional genomics in the era of next generation sequencing: An update, The Crop Journal (2017), https://doi.org/10.1016/j.cj.2017.09.003