

# Accessing complex crop genomes with next-generation sequencing

David Edwards · Jacqueline Batley ·  
Rod J. Snowdon

Received: 31 January 2012 / Accepted: 8 August 2012 / Published online: 5 September 2012  
© Springer-Verlag 2012

**Abstract** Many important crop species have genomes originating from ancestral or recent polyploidisation events. Multiple homoeologous gene copies, chromosomal rearrangements and amplification of repetitive DNA within large and complex crop genomes can considerably complicate genome analysis and gene discovery by conventional, forward genetics approaches. On the other hand, ongoing technological advances in molecular genetics and genomics today offer unprecedented opportunities to analyse and access even more recalcitrant genomes. In this review, we describe next-generation sequencing and data analysis techniques that vastly improve our ability to dissect and mine genomes for causal genes underlying key traits and allelic variation of interest to breeders. We focus primarily on wheat and oilseed rape, two leading examples of major polyploid crop genomes whose size or complexity present different, significant challenges. In both cases, the latest DNA sequencing technologies, applied using quite different approaches, have enabled considerable progress

towards unravelling the respective genomes. Our ability to discover the extent and distribution of genetic diversity in crop gene pools, and its relationship to yield and quality-related traits, is swiftly gathering momentum as DNA sequencing and the bioinformatic tools to deal with growing quantities of genomic data continue to develop. In the coming decade, genomic and transcriptomic sequencing, discovery and high-throughput screening of single nucleotide polymorphisms, presence–absence variations and other structural chromosomal variants in diverse germplasm collections will give detailed insight into the origins, domestication and available trait-relevant variation of polyploid crops, in the process facilitating novel approaches and possibilities for genomics-assisted breeding.

## Introduction

Polyploidy is a key evolutionary mechanism that has been a driving factor in the success of some of the most important domesticated plants (Leitch and Leitch 2008). Hexaploid wheat (*Triticum aestivum*) and oilseed rape/canola (*Brassica napus*) are prime examples of globally important cereal and oilseed plants with large, complex genomes. The bread wheat genome is allohexaploid, consisting of six sets of chromosomes originating from three distinct diploid genomes (termed A, B and D). The three diploid donor species each contributed seven pairs of chromosomes, resulting in 21 chromosome pairs. The donor species diverged between 2.5 and 6 MYA, and two distinct interspecies hybridisation events occurred which serially gave rise to the current polyploid bread wheat varieties. The first event, between 0.5 and 3 MYA, was a hybridisation between *T. urartu* ( $A^uA^u$ ) and an unknown

---

Communicated by R. Varshney.

D. Edwards  
Australian Centre for Plant Functional Genomics, School  
of Agriculture and Food Sciences, University of Queensland,  
Brisbane, QLD 4072, Australia

J. Batley (✉)  
Centre of Excellence for Integrative Legume Research, School  
of Agriculture and Food Sciences, University of Queensland,  
Brisbane, QLD 4072, Australia  
e-mail: j.batley@uq.edu.au

R. J. Snowdon  
Department of Plant Breeding, IFZ Research Centre for  
Biosystems, Land Use and Nutrition, Justus Liebig University,  
Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany