

Wheat genome structure and function: genome sequence data and the International Wheat Genome Sequencing Consortium.

P. Moolhuijzen^A, *D. S. Dunn*^A, *M. Bellgard*^A, *M. Carter*^B, *J. Jia*^C, *X. Kong*^C, *B. S. Gill*^D, *C. Feuillet*^E, *J. Breen*^A and *R. Appels*^{A, F}

^A Centre for Comparative Genomics, Murdoch University, Murdoch, WA 6150, Australia.

^B State Agric Biotechnology Centre, Murdoch University, Murdoch, WA 6150, Australia.

^C Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, 100081, Beijing, China.

^D Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA.

^E UMR ASP 1095, INRA, University Blaise Pascal, 63100 Clermont-Ferrand, France.

^F Corresponding author. Email: rappels@cgc.murdoch.edu.au

This paper is dedicated to Professor Bob McIntosh, in recognition of his tireless efforts to critically analyse the work carried out on the genetics of wheat.

Abstract

Genome sequencing and the associated bioinformatics is now a widely accepted research tool for accelerating genetic research and the analysis of genome structure and function of wheat because it leverages similar work from other crops and plants. The International Wheat Genome Sequencing Consortium addresses the challenge of wheat genome structure and function and builds on the research efforts of Professor Bob McIntosh in the genetics of wheat. Currently, expressed sequence tags (ESTs; ~500 000 to date) are the largest sequence resource for wheat genome analyses. It is estimated that the gene coverage of the wheat EST collection is ~60%, close to that of *Arabidopsis*, indicating that ~40% of wheat genes are not represented in EST collections. The physical map of the D-genome donor species *Aegilops tauschii* is under construction (<http://wheat.pw.usda.gov/PhysicalMapping>). The technologies developed in this analysis of the D genome provide a good model for the approach to the entire wheat genome, namely compiling BAC contigs, assigning these BAC contigs to addresses in a high resolution genetic map, filling in gaps to obtain the entire physical length of a chromosome, and then large-scale sequencing.