

GENOMICS

Sequence resources for crucial crops



Photodisc/Getty

Genomic resources will be essential for crop plant improvement, but several crops have large, repeat-rich genomes that present major technical challenges. Therefore, reference genomes are not yet available for some of the most nutritionally important species. Nevertheless, recent work on the barley and bread wheat genomes has used sophisticated approaches to enhance greatly the resources that could be used for genome-assisted research and development.

The International Barley Genome Sequencing Consortium has developed a multi-layer resource that integrates a genome-wide physical map, genome sequence data and expression data. The physical map was generated through a bacterial artificial chromosome (BAC) contig assembly and covers >95% of the 5.1 Gb barley haploid genome. Shotgun sequencing of some of the BAC clones — focusing on those containing genes — provided sequence data that could be fitted onto the map framework. This sequence-enriched map facilitated the integration of short-read whole-genome shotgun sequence data from genomic DNA. Despite the problems presented by the high levels of repeat sequences — ~84% of the genome, as assessed by this study

— this work has provided a valuable high-resolution genetic map of ~4 Gb of the genome.

The authors added important functional information to the barley sequence assembly through high-throughput RNA sequencing (RNA-seq) of samples from eight stages of development and through sequencing of full-length cDNA sequences. As well as annotating ~26,000 'high-confidence' genes, this work provided intriguing insights into the regulation of gene expression in this crop. For example, the authors found evidence of extensive regulation through alternative splicing linked to nonsense-mediated decay. In addition, the authors made comparisons with four other cultivars (the sequence assembly used the 'Morex' cultivar) to characterize genome diversity. This enabled a single-nucleotide variant map to be generated that will aid breeding efforts.

The difficulties of the barley genome might seem small when compared with the ~17 Gb hexaploid genome of bread wheat, which also has an estimated repeat content of ~80%. Brenchley *et al.* have now created an assembly of the gene content of bread wheat by 454 pyrosequencing. The authors harnessed the known genome sequences of three other grasses — namely, rice, sorghum and *Brachypodium distachyon* — to assemble groups of orthologous gene sequences. This orthologous group assembly leads to an estimate of a total of 94,000–96,000 genes in the wheat genome.

The hexaploid genome originates from three diploid progenitor genomes designated AA, BB and DD, and the authors made comparisons with the genomes of relatives of these ancestors of wheat and assigned most wheat genes to one of the progenitor genomes. This has provided a resource of >132,000 SNPs that can be used in future quantitative trait studies of agricultural relevance. Insights into changes in gene content linked to domestication were also uncovered. For example, comparisons with the diploid goat grass *Aegilops tauschii*, which is known to be the donor of the DD genome, suggested that there may have been increases in gene families associated with nutrition and energy metabolism in the wheat lineage.

Although the ultimate goal remains a high-quality reference genome, these examples show that innovative strategies that draw on different sequencing, mapping and comparative genomics approaches can provide much needed resources to accelerate progress in agricultural research.

Mary Muers

ORIGINAL RESEARCH PAPERS The International Barley Genome Sequencing Consortium. A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491**, 711–716 (2012) | Brenchley, R. *et al.* Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* **491**, 705–710 (2012)
FURTHER READING Morrell, P. A., Buckler E. S. & Ross-Ibarra, J. Crop genomics: advances and applications. *Nature Rev. Genet.* **13**, 85–96 (2012)