Comparative genomics as a guide to the wheat genome

<u>Bevan M</u>[1], Febrer M[1], Wright J[1], Yang A[1], McKenzie N[1], Salse J[2] and Feuillet C[2]
[1] John Innes Centre, Colney Lane, Norwich NR4 7UJ, UK
[2]INRA UMR INRA-UBP 1095, 63100 Clermont-Ferrand , France

The economic importance of wheat as a primary food and feed source demands the rapid development of genomic resources to support breeding and gene discovery. However, the hexaploid wheat genome presents a unique challenge to genome scientists due to its large size and complexity. The International Wheat Genome Sequencing Consortium aims to exploit chromosome – sorting methods to construct physical maps of BACs that are anchored to high density genetic maps. Recent progress in chromosome 3B has established this as a feasible genomics strategy, which is now being applied systematically to most chromosomes. Concurrently major advances in sequence technology have dramatically reduced the costs of sequence generation. However, a BAC- based genome sequencing strategy is limited by the rate at which BAC contigs can be aligned to the genetic map, the extent of coverage of the chromosome, and the accuracy of placement and orientation of BACs. On the other hand, whole genome shotgun sequencing and *de novo* assembly of the wheat genome using next generation sequencing may be confounded by repeat complexity, complex polyploid gene relationships, and the sheer scale of the computational task.

We have initiated a comparative genomics strategy to achieve a successively more detailed and accurate sequence of wheat genes and their correct chromosomal locations. The following steps are envisaged:

- Sequencing the genome of *Brachypodium distachyon* and establishing high quality annotations of genes and repeats. This forms a foundation for comparative genomics in the Pooideae.
- Aligning the wheat and *Brachypodium* genomes according to bin- mapped wheat EST sequences.
- Sequencing the wheat transcriptome to deep coverage using next generation sequencing and aligning the transcriptome assemblies to *Brachypodium* gene models. This provides the first approximation of gene order in wheat, and facilitates bioniformatics analyses for establishing gene family and homoeologous relationships, and for designing COS markers.
- Devising COS markers from the wheat transcriptome assemblies for aligning the genetic map of wheat to the *Brachypodium* genome sequence assemblies to a high resolution.
- Defining segments of conserved gene order between wheat and *Brachypodium* from which genic probes can be made for anchoring BAC MTPs (minimal tiling paths).
- Extending and contiguating BAC MTPs using conserved gene order.
- Using the assembled BES (BAC end sequences) and the wheat transcriptome sequences to facilitate assembly of wheat genome sequences in their correct chromosomal location.