

Meeting report

Applying plant genomics to crop improvement

Michael Bevan* and Robbie Waugh[†]

Addresses: *John Innes Centre, Colney Lane, Norwich NR4 7UH, UK. [†]Scottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, UK.

Correspondence: Michael Bevan. E-mail: michael.bevan@bbsrc.ac.uk

Published: 15 February 2007

Genome Biology 2007, **8**:302 (doi:10.1186/gb-2007-8-2-302)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2007/8/2/302>

© 2007 BioMed Central Ltd

A report of the European Science Foundation-Wellcome Trust Conference on Crop Genomics, Trait Analysis and Breeding, Hinxton, UK, 8-11 November 2006.

Nearly 150 scientists and industrialists attended a recent meeting outside Cambridge to review progress in the application of genomics to crop plant improvement. The meeting covered a wide range of topics, from genome sequencing methods to marker-assisted breeding for wheat improvement. In her opening address, Julia Goodfellow (Biotechnological and Biological Sciences Research Council, Swindon, UK) described the increasing need for more healthy and nutritious food produced in environmentally sustainable ways and the need to translate the fruits of basic research in model species into crop improvement. The meeting established that such a research pipeline is a high priority and that genomics provides the means to achieve it.

Crop plants often have large and complex genomes; the maize genome, for example, is around 2.5 gigabase pairs (10^9 base pairs), approximately the same size as that of humans. Richard McCombie (Cold Spring Harbor Laboratory, New York, USA) described the remarkable progress being made in maize genome sequencing at Washington University in St Louis by Richard Wilson and colleagues. Their strategy is to sequence 19,000 bacterial artificial chromosome clones (BACs) from the minimal tiling path - the minimum number of overlapping clones that cover the whole genome - to finish the sequence, and to annotate it. One year into the project, 30% of the clones are in the shotgun-sequencing phase and nearly 10% have been finished using capillary-based sequencing.

McCombie noted that new sequencing technologies could be applied to BAC-based strategies so as to dramatically increase the speed and cost-effectiveness of genome sequencing. This

topic was addressed by David Bentley (Solexa, Great Chesterford, UK), who described recent progress in developing high-throughput chip-based sequencing technologies that have already radically improved the throughput and cost-effectiveness of sequencing. Using simultaneous base-by-base sequencing of up to 40 million immobilized templates, average read lengths of 35 bases of high-quality data can be generated, within a few hours. Improvements in dye chemistry, polymerase and imaging methods promise to increase read lengths, but already the huge depth of coverage can give accuracies of over 99%. The range of applications for this technology, apart from resequencing, includes digital gene-expression profiling, BAC sequencing and polymorphism detection, all of which are currently very cost-effective using Solexa technology. On cue, a project that could really only be countenanced with such breakthrough technologies was described by Catherine Feuillet (French National Institute for Agricultural Research (INRA), Clermond-Ferrand, France). She presented exciting progress in tackling the Mount Everest of genomes - the 16 gigabase-pair hexaploid genome of bread wheat. Using flow-sorted chromosomes and chromosome arms, chromosome-specific BAC libraries have been made for physical mapping. The challenge now is to apply this approach to all 21 chromosomes, and then to deploy multiple approaches for aligning BACs. Regions of interest could then be sequenced by Solexa's or related methods.

Access to high-throughput genome sequencing methods also underlies the ambition to sequence no fewer than 100 solanaceous species as part of the international SOL project, as reported by Rene Klein Langhorst (Centre for Biosystems Genomics, Wageningen, The Netherlands). This group of plants accounts for a huge range of economically important species, including tobacco, coffee, many vegetables and a vast range of untapped plants. Ian Bancroft (John Innes Centre, Norwich, UK) described the international effort to sequence a *Brassica* genome, and some initial findings from

comparative genomics. A collaboration involving Korea, Australia, China and the UK will sequence *Brassica rapa* (the turnip and Chinese cabbage family of crops). This will also throw some light on the genome of *B. napus* (oilseed rape) as the AA genome of *B. napus* (which is a tetraploid denoted AACC) comes from *B. rapa*. Despite their close evolutionary relationship to the reference *Arabidopsis thaliana* genome, *Brassica* genomes are exceptionally complex, as two rounds of genome duplication since their last common ancestor with *Arabidopsis*, followed by diploidization and hybridization events, have led to extensive gene loss, chromosome rearrangements and additional segmental duplications. The *B. rapa* genome sequence promises to reveal both agronomically important genes and a detailed record of chromosome dynamics wrought by evolution and human intervention.

Michele Morgante (University of Udine, Udine, Italy) expanded on the discussion of genome dynamics by introducing the concepts of the 'pan-genome' and the 'core-genome', based on his analysis of the maize genome - or rather genomes. He has found a remarkable plasticity in the maize genome that is probably typical of most genomes with a high repeat content. Maize appears to be hypermutable, mainly as a result of retrotransposon activity, with one polymorphic insertion or deletion every 2 kilobases between the commercial cultivars Mo17 and W73. The pan-genome encompasses the spectrum of diversity overlaid on a set of core, conserved components, and understanding this hypervariation promises to be a major outcome of the sequencing of maize and the resequencing of cultivars. Again, high-throughput genome sequencing methods will play a central role.

The epigenome of *A. thaliana*, comprising the reversible modifications to chromatin that affect gene expression, is also highly polymorphic, according to Robert Martienssen (Cold Spring Harbor Laboratory). Epigenetic marks in both heterochromatin and genes were found to be hypervariable between the ecotypes Columbia and Landsberg erecta, but this did not appear to affect gene expression.

The applications of DNA markers to identifying and selecting wheat varieties were described by both Simon Berry and Bill Angus (Nickerson-Advanta, Lincoln, UK). Many examples of the benefits of DNA marker technology in crop breeding were described, including tracking desired loci or genes, pyramiding genes in robust combinations and increased efficiency in selecting complex traits. Markers for tracking multiple traits, such as disease tolerance, lodging resistance, height and grain quality are being used to achieve year-on-year improvements in yield of nearly 2% a year.

The need for continuous improvement and adaptability of crops was put into stark perspective by Steve Jefferies (Australian Grain Technologies, Adelaide, Australia). In the worst drought in over 100 years, more than 25% of

Australian grain production has been lost. An aggressive breeding strategy is under way to make better-adapted varieties using marker-assisted breeding. The speed and precision of breeding is steadily increasing, and the availability of physical maps and genome sequences will initiate another step-change towards so-called 'molecular breeding' strategies.

Combating global climate change and securing sufficient energy supplies are now major topics in which plant research has a centrally important role. The exciting promise of bioenergy crops was revealed by Angela Karp (Rothamsted Research, Harpenden, UK). She described progress in improving *Salix* (willow) for bioenergy production. DNA-based technologies promise to dramatically reduce the long breeding cycle of this dioecious genus (one in which there are separate male and female trees). Markers associated with useful traits such as branching, stem diameter and beetle resistance have already been identified, and excellent progress in yield improvements on marginal land has been made during the first rounds of breeding.

The talks at the meeting showed that crop plant genomics is now firmly embedded in the breeding programs of many crops, and major gains are being achieved while accelerating the rate and precision of breeding and enabling the assembly of multiple traits in new elite lines. Technologies and funding opportunities are now available for generating and assembling the genome sequences for many of our major crop plants for food and energy production. This will take breeding and trait analysis to a new level of precision and productivity that will be needed to meet the major challenge of food and energy security in the coming years.