

Meeting report

Discovering the seeds of diversity in plant genomes

James A Birchler and Kathleen J Newton

Address: Department of Biological Sciences, University of Missouri, Columbia, MO 65211, USA.

Correspondence: James A Birchler. E-mail: BirchlerJ@Missouri.edu

Published: 26 April 2004

Genome Biology 2004, **5**:323

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2004/5/5/323>

© 2004 BioMed Central Ltd

A report on the Keystone Symposium 'Comparative Genomics of Plants', Taos, USA, 4-9 March 2004.

This meeting, organized by Richard Flavell (Ceres, Malibu, USA) and Rob Martienssen (Cold Spring Harbor Laboratory, USA), brought together a diverse group of speakers for a discussion of plant genome organization and the types of variation that exist between and within species. This report focuses on the consequences of this variation on phenotype, and the basis of this variation at the DNA and epigenetic levels.

Keynote speaker Steve Tanksley (Cornell University, Ithaca, USA) provided a historical account of the molecular identification of quantitative trait loci (QTLs). Those QTLs for which a molecular basis have been defined in his laboratory, using tomato as a model system, point to that fact that most such loci are regulatory in nature. This conclusion was bolstered in the talk by David Jackson (Cold Spring Harbor Laboratory, USA), who reported that *ramosai* in maize, a gene involved in floral branching, has been identified as a transcription factor. This gene has also been shown to be a QTL candidate gene in mapping studies of populations that have different numbers of tassel branches. In addition, the maize *fasciated ear2* (*fae2*) gene involved in determining kernel-row number is homologous to the *CLAVATA2* floral development regulator in *Arabidopsis*, and QTL analysis also shows an effect on floral development at this locus in maize. Ed Buckler (Cornell University, Ithaca, USA) reported that the maize *Dwarf8* gene, which encodes a transcription factor, is associated with a QTL involved in the control of plant height and flowering time. These accumulating data support the generalization introduced by Tanksley that genetic variation that leads to changes in plant form tends to affect gene regulation, and that one should focus the search for evolutionarily important variation on regulatory

changes, particularly those that fall outside of amino-acid-coding regions. Indeed, Stephen Goff (Syngenta, San Diego, USA) has taken a genomics approach to identify *cis*-regulatory sites involved with MYB transcription-factor activity in rice, and the network of genes controlled by these regulators. Comparison with maize indicates that these promoters show conservation across the cereals; furthermore, the regulatory regions could be swapped across species in transgenic plants and still produce similar tissue specificities.

Further discussion of regulatory dosage effects and evolution came in the talk by Michael Freeling (University of California, Berkeley, USA). He analyzed the retention of genes in *Arabidopsis* following a very ancient allotetraploidization event. Most of the duplicate genes have been deleted since this event, thus returning most of the genome to the diploid level. Interestingly, there was a strong tendency for transcription factors to be retained in the duplicated state. Freeling's discussion of why this should be the case centered on the potential of regulatory genes to evolve new functions and hence be retained in the genome. Other potential explanations discussed included the need to maintain a balance of interacting regulators and the fact that the deletion of individual transcription-factor loci would mimic haplo-insufficiency and trigger selection to maintain the regulatory balance. Tom Osborn (University of Wisconsin, Madison, USA) noted that polyploidy increases the variation in dosage-regulated gene expression, as determined using the Flowering Locus C (FLC) in *Arabidopsis* and *Brassica* as model systems. This gene, first defined in *Arabidopsis*, affects flowering time, having a delaying effect with increasing dosage. In the polyploid relatives of *Arabidopsis* from the genus *Brassica*, greater variation in the quantity of the FLC gene product is possible, and this can be used to generate a much larger span in flowering time in the polyploids. The take-home message is that gene-regulatory networks are a major contributor to morphological and quantitative trait variation.

Another common theme was the variation found in plant genomes. Buckler reported that genome variation within maize is greater than in humans. By documenting this extensive variation and the phenotypes of multiple inbred lines, association analysis has the potential to identify the nucleotide polymorphisms that are responsible for the phenotypic differences. In contrast, the level of variation in the rice genome is lower than that of maize, as was revealed in a talk about sequencing the rice genome from Takuji Sasaki (National Institute of Agrobiological Sciences, Japan), and one about the genomic analysis of different rice species from Susan McCouch (Cornell University, Ithaca, USA). Another type of variation in plant genomes was discussed by Scott Tingey (Dupont, Newark, USA). Bacterial artificial chromosome (BAC) contigs were compared between two common inbred lines in maize, namely B73 and Mo17. The arrangement of genes in homologous regions is dramatically different in these two lines - not only are transposable elements variable, but the genes show different arrangements. Within one region tested, rice probes applied to maize indicated a presence in both B73 and Mo17 for only slightly more than half of the genes examined. In some cases the genes have been lost entirely, but in other cases they reside elsewhere in the genome. These results indicate that diversity is caused by mechanisms other than point mutation.

One of us (J.B.) reported the development of a maize karyotyping method that relies on fluorescent *in situ* hybridization of tandemly repetitive sequence clusters, such as centromere repeats, ribosomal RNA genes, knob heterochromatin and subtelomeric sequences. These cytological features show a large variation in quantity across various maize inbred lines as well as in their presence at any one site in the genome. Mechanisms that could generate such diversity involve the action of transposable elements. Katrien Devos (University of Georgia, Athens, USA) discussed the expansion of retroelements in the grass family. Most transposable elements are relatively young and hence can account for many genomic differences among species. The sequences of such elements degrade over evolutionary time via point mutations, but also by unequal recombination to produce solo LTRs (long terminal repeat elements); illegitimate recombination can also degrade the single LTRs. Bursts of transposition of these elements produce extensive variation in genome size within the plant kingdom, as was noted by Ilia Leitch (Royal Botanic Garden, Kew, UK). Sue Wessler (University of Georgia, Athens, USA) reported an analysis of transposable elements in the rice genome. There are three major types of transposable element in this species. The small MITES (miniature inverted repeat transposable elements) are present as 85,000 copies of various families and are typically associated with genes; the large helitrons are present in 10,000 copies, and the various forms of retroelements total about 6,000 copies. In contrast to MITES, retroelements tend to cluster in pericentric regions. Mutator-like elements, called PackMULES, have captured host genes at the DNA

level and mobilize these sequences throughout the genome. To date, approximately 3,500 PackMULES have been recognized in rice. The ability of PackMULES to mobilize genes provides one mechanism to generate the diversity of genome arrangements described above.

Variation that does not depend on changes in DNA sequence involves epigenetic modifications of chromatin and DNA. Luca Comai (University of Washington, Seattle, USA) reported the changes that are associated with allopolyploidy and autopolyploidy in *Arabidopsis* and its close relatives. Some genes are silenced in newly formed allopolyploids, leading to the potential for new phenotypes. Comai's group generated new autopolyploids from different ecotypes, and showed that they react differently in interploidy crosses. In some cases, the seeds are viable in such crosses in one direction of cross but not the other. In other combinations of interploidy crosses of different ecotypes, the seed failure occurs in both directions of reciprocal crosses. An interesting set of recombinant inbred lines was generated from a cross of tetraploids of one ecotype and diploids of another. The re-establishment of tetraploids from inbreeding the triploid F1 generation suggested the presence of a gene that fosters the reformation of tetraploids. Craig Pikaard (Washington University, St Louis, USA) described the phenomenon of nucleolar dominance, where the ribosomal RNA genes of one parent are expressed in a hybrid or allopolyploid, while those of another parent are repressed. In the allotetraploid *Arabidopsis suecica*, formed from the genomes of *Arabidopsis thaliana* and *Arabidopsis arenosa*, the *A. thaliana* genes are silenced. The active genes are associated with modifications of histone H3, namely the methylation of the lysine at position 4, whereas the inactive genes are associated with H3 methylated at the lysine at position 9. These modifications are correlated with DNA hypomethylation and DNA hypermethylation, respectively. Inactivation of two histone-deacetylase genes using RNA interference (RNAi) reactivates the silent gene copies and results in a loss of DNA methylation as well.

Martienssen reported the analysis of a region of heterochromatin in *Arabidopsis*. A chromosomal-tiling path was created in 1 kilobase segments that spanned a 1.5 megabase region. An analysis of the spectrum of methylation of H3 at lysine 9 and DNA shows a good correlation with the placement of small interfering RNA (siRNA) origins on the genomic sequence. This correlation suggests that siRNAs might guide the modification of histone H3 and DNA in *Arabidopsis*. Interestingly, mutations in the *Dicerlike* or *Argonaute* genes, whose products are thought to participate in the RNAi pathway, do not affect the expression of most transposable elements, whereas the *Decrease in DNA Methylation1* (*DDM1*) gene does affect their expression. Redundancy of *Argonaute* genes in a complex family within the *Arabidopsis* genome may explain this result. Eric Richards (Washington University, St Louis, USA) described

the accumulation of epi-mutations in the *ddm1* mutant background. These results raise the possibility that epigenetic variation might exist in natural populations and might play a role in plant evolution. DNA methylation is also associated with imprinted genes in the endosperm of plants, a topic that was discussed by Robert Fischer (University of California, Berkeley, USA). The *Medea* gene is expressed in the female gametophyte and from the maternal alleles in the endosperm, but is not expressed from the paternal contribution to the endosperm. Release of gene silencing on the maternally inherited *MEDEA* allele requires the DNA glycosylase DEMETER (DME). It is thought that the *Demeter* gene product nicks the methylated DNA in the promoter of the *Medea* gene; the nick is then repaired, thus removing the methylated epigenetic mark on the maternal allele. In summary, the meeting made clear that comparative analyses free researchers from species constraints and allow the elucidation of the roles that processes such as polyploidy, transposon activation, epigenetic modification and altered gene regulation play in the dynamics of genome evolution in plants.

Acknowledgements

We thank Eric Richards for comments and suggestions on the manuscript.