

- Simillion C, Vandepoele K, Saeys Y, Van de Peer Y. 2004. Building genomic profiles for uncovering segmental homology in the twilight zone. *Genome Research* 14: 1095–1106.
- Simillion C, Vandepoele K, Van Montagu MCE, Zabeau M, Van de Peer Y. 2002. The hidden duplication past of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* 99: 13627–13632.
- Stebbins G. 1966. Chromosomal variation and evolution; polyploidy and chromosome size and number shed light on evolutionary processes in higher plants. *Science* 152: 1463–1469.
- Vandepoele K, Simillion C, Van de Peer Y. 2003. Evidence that rice and other cereals are ancient aneuploids. *Plant Cell* 15: 2192–2202.
- Vision T, Brown D, Tanksley S. 2000. The origins of genomic duplications in *Arabidopsis*. *Science* 290: 2114–2117.
- Wang X, Shi X, Hao BL, Ge S, Luo J. 2005. Duplication and DNA segmental loss in rice genome and their implications for diploidization. *New Phytologist* 165: 937–946.
- Yu J, Hu SN, Wang J, Wong GKS, Li SG, Liu B, Deng YJ, Dai L, Zhou Y, Zhang XQ, Cao ML, Liu J, Sun JD, Tang JB, Chen YJ, Huang XB, Lin WYeC, Tong W, Cong LJ, Geng JN, Han YJ, Li L, Li W, Hu GQ, Huang XG, Li WJ, Li J, Liu ZW, Liu JP, Qi QH, Liu JS, Li T, Wang XG, Lu H, Wu TT, Zhu M, Ni PX, Han H, Dong W, Ren XY, Feng XL, Cui P, Li XR, Wang H, Xu X, Zhai WX, Xu Z, Zhang JS, He SJ, Zhang JG, Xu JC, Zhang KL, Zheng XW, Dong JH, Zeng WY, Tao LYeJ, Tan J, Ren XD, Chen XW, He J, Liu DF, Tian W, Tian CG, Xia HG, Bao QY, Li G, Gao H, Cao T, Zhao WM, Li P, Chen W, Wang XD, Zhang Y, Hu JF, Liu S, Yang J, Zhang GY, Xiong YQ, Li ZJ, Mao L, Zhou CS, Zhu Z, Chen RS, Hao BL, Zheng WM, Chen SY, Guo W, Li GJ, Liu SQ, Tao M, Zhu LH, Yuan LP, Yang HM. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). *Science* 296: 79–92.

Key words: ancient duplication, *Arabidopsis*, cereal crops, *Oryza* (rice), polyploidization.

Meetings

A pioneer perspective on adaptation

Functional genomics of environmental adaptation in *Populus*: the 12th *New Phytologist* Symposium, Gatlinburg, TN, USA, October 2004

In its natural habitat, the black cottonwood tree (*Populus trichocarpa* Hooker) is a pioneer species that thrives in the dynamic, resource-rich environments created by massive flooding (Fig. 1; Braatne *et al.*, 1996). With the recent public release of the complete genome sequence of *P. trichocarpa* (<http://genome.jgi-psf.org/Poptr1/Poptr1.home.html>), an analogous flood of data is creating unprecedented opportunities in basic and applied research on this model tree (Strauss & Martin, 2004; Tuskan *et al.*, 2004). This is a landmark event for ecological research in particular, because unlike herbaceous model plants like *Arabidopsis*, corn, and rice, *Populus* is a wild, perennial plant that constitutes a dominant component of many ecosystems throughout the northern hemisphere (Braatne *et al.*, 1996). The genome sequence provides a cryptic blueprint of the molecular underpinnings of adaptation in natural populations, raising the possibility of linking molecular polymorphisms to adaptively significant phenotypic variation, a feat that has been accomplished only rarely in plant research until recently (Remington *et al.*, 2001; Jackson *et al.*, 2002). Breakthroughs in the understanding of ecologically significant molecular variation will require continued technological advancement and investment in genomic resources for additional ecologically

important species, as well as cross-disciplinary collaboration in ecology, population genetics and molecular biology (Feder & Mitchell-Olds, 2003). To facilitate this collaboration, the 12th *New Phytologist* symposium brought together a diverse group of ecologists, geneticists and molecular biologists to explore the prospects for determining molecular determinants of adaptive variation in *Populus* and other species. Rapid progress is likely in the understanding of molecular underpinnings of adaptation in *Populus* due to a rapidly expanding molecular toolbox, an energized and growing research community, and ecological assets that are unparalleled among sequenced model plants.

Adaptation research in forest trees

Adaptation research in forest trees has a long history, driven in part by the commercial importance of producing well-adapted and improved seed stocks for reforestation, and in part by the ecological importance of forest trees (Howe *et al.*, 2003). Initially, studies focused exclusively on adaptive traits, and usually involved planting diverse provenances in a common garden and performing a battery of measurements over a long time span. Such studies continue to be important because they provide the most direct means for determining the distribution of adaptive variation on the landscape, the degree of genetic and environmental control of a wide range of complex traits, and the extent of genotype by environment interactions (Zobel & Talbert, 1984). However, such studies provide little insight into the underlying genetic architecture of complex traits (i.e. the number and strength of genes involved, and the degrees of additivity and dominance of individual genetic loci).



Fig. 1 A *Populus trichocarpa* stand on the Willamette River near Corvallis, OR, USA.

The introduction of biochemical and molecular markers revolutionized the study of population genetics and adaptive variation. Putatively neutral molecular markers have allowed delineation of the genetic structure of natural tree populations and inferences about the evolutionary history of species, and levels of historical and contemporary gene flow (Mitton, 1994; Smouse & Sork, 2004), all of which contribute to the understanding of adaptive variation. Furthermore, comparison of genetic population structure at neutral loci (F_{ST}) to the structure of quantitative traits (Q_{ST}) allows inferences about selective forces acting on populations (Howe *et al.*, 2003). Finally, neutral markers have allowed determination of the genetic architecture of complex traits by exploiting linkage to quantitative trait loci (QTL) in structured pedigrees (Lynch & Walsh, 1998), which raised the possibility of identifying the molecular determinants of phenotypic variation (Remington *et al.*, 2001).

Associating genotypes with phenotypes

The current feasibility of studying molecular adaptive variation is due to rapid advances in marker technology and the accumulation of a critical mass of structural and functional genomics data from model organisms. Optimal strategies for identifying and characterizing adaptively significant molecular polymorphisms depend upon the details of the organism, traits and populations to be studied. One key consideration is the amount and organization of genetic diversity in the species. Ideally, the organism should harbor a large amount of diversity due to a long history of large effective population

sizes, as seems to be the case for many forest trees (Neale & Savolainen, 2004), and/or a high inherent mutation rate. Another important consideration is the prevailing level of linkage disequilibrium (LD) in the study population, because this determines the size of genomic regions that are marked by individual polymorphic genetic markers. A high level of LD is desirable to allow efficient scanning of the genome for marker–phenotype associations and rapid determination of the genetic architecture of a complex trait. For example, initial QTL analyses are usually conducted with early-generation pedigrees derived from divergent parents, because the level of LD is quite high in this situation due to a limited number of opportunities for recombination during meiosis. However, a major disadvantage of populations with high levels of LD is that many genes and polymorphisms are contained in the chromosomal region represented by the marker, thus greatly complicating the identification of causative polymorphisms. In contrast, populations with low levels of LD, including wild populations of many tree species (Neale & Savolainen, 2004; Yin *et al.*, 2004), potentially allow the identification of very small chromosomal segments that are associated with traits of interest, bringing researchers much closer to the ultimate goal of identifying the actual molecular polymorphisms that are responsible for adaptive phenotypes. The trade-off is that very large numbers of markers are required to scan the entire genome for associations with complex traits. For example, the level of LD appears to decline within about 1.5 kb in loblolly pine genome (Brown *et al.*, 2004), meaning that at a minimum at least one marker per gene would be required for

whole-genome association scans, and many more if putative noncoding regions are to be surveyed as well. At the present time, this is an unfeasible level of genotyping at the levels of funding typically available for plant genetics research.

A common solution to this conundrum is the candidate gene approach, in which genes are identified *a priori* as likely candidates for the trait of interest and variation in those genes is screened in the test population. This approach has had some success in forest trees, as described at the meeting by David Neale (USDA Forest Service, Davis, CA, USA) and Christophe Plomion (INRA, Cestas, France) for pine, Douglas-fir and oak. Associations have been detected between candidate gene polymorphisms and wood quality, drought, disease resistance and phenological traits, despite the fact that a relatively small number of candidate genes have been screened thus far.

Despite these initial successes, the candidate gene approach is widely acknowledged to contain some significant flaws. One of the principal problems is that the process of selecting candidates limits possibilities for novel gene discovery. The most common source of candidates is information from model herbaceous organisms, principally *Arabidopsis*. Whereas there is clearly substantial conservation of gene structure and function between *Arabidopsis* and trees, even including the distantly related gymnosperms (Kirst *et al.*, 2003), there is also substantial divergence reflected by the stark differences in morphology and life history. Therefore, many genes and polymorphisms that are important in tree adaptation will have to be discovered in trees directly. Furthermore, candidate gene approaches may provide a limited view of the genetic architecture of traits and may fail to capture complex patterns of epistasis. For example, Michael Purugganan (North Carolina State University, Raleigh, NC, USA) described a situation in *Arabidopsis* where a haplotype of *Flowering Locus C* was associated with latitudinal variation in flowering time only when an active allele of the *Frigida* gene was present. This association and interaction would only be detected with some *a priori* knowledge of the pathway, which could be quite different in a perennial tree like *Populus* or a distantly related gymnosperm like pine.

Candidate genes in *Populus*

Fortunately, there are substantial functional genomics efforts in forest trees, especially in *Populus* (Bhalerao *et al.* 2003), and these are generating tree-specific candidate genes for a host of important traits. Research described at the meeting was representative of worldwide efforts, which typically focus on traits that are tree-specific, including dormancy, maturation and wood-formation (Brunner & Nilsson, 2004; Pilate *et al.*, 2004), as well as studies of tree-specific symbionts, pathogens and insect pests (Arimura *et al.*, 2004; Grunze *et al.*, 2004; Martin *et al.*, 2004a). In addition to intensive studies of single genes and gene families, there are also substantial efforts to perform broad transcript profiling using expressed sequence tag (EST) sequencing of

libraries prepared from a wide variety of tissues and treatments, and microarray analyses (Smith *et al.*, 2004; Sterky *et al.*, 2004). Several laboratories have prepared large cDNA arrays containing up to 50% of the predicted genes in the *Populus* genome, and soon whole-genome oligonucleotide microarrays will be available for *Populus* through NimbleGen (Madison, WI, USA) and Affymetrix (Santa Clara, CA, USA). Furthermore, several groups are developing high-throughput proteomic and metabolomic platforms for *Populus*, raising the possibility for metabolic pathway modeling. Finally, one of the primary strengths of *Populus* as a model species is the relative ease with which it can be genetically transformed and propagated. This has facilitated 'forward-genetics' gene discovery approaches using activation tagging and enhancer traps, as well as 'reverse-genetics' through the creation of gene knockouts and overexpression. Transformation provides an excellent complement to studies in natural population because strong, dominant mutations can be induced that would be unlikely to persist in natural populations due to selective sweeps (Brunner *et al.*, 2004). Large collections of informative transgenic lines are being generated by *Populus* researchers, and efforts are underway to develop centralized germplasm repositories and distribution systems (Tsai & Hubscher, 2004). Taken together, all of these approaches afford the opportunity to generate extensive lists of candidate genes that can be tested in association studies in natural populations. These efforts are described in greater detail in the International *Populus* Genome Consortium science plan (http://www.ornl.gov/sci/ipgc/the_populus_genome_science_plan.pdf)

Even with extensive lists of genes and high-throughput genotyping capabilities in hand, candidate gene association studies may still only provide a partial picture of the molecular underpinnings of adaptation. If surveys focus on coding regions exclusively, *cis*- and (especially) *trans*-acting regulatory elements may be missed. Furthermore, as was pointed out in multiple presentations at the meeting, many gene functions are not apparent under controlled laboratory conditions, and experiments in realistic field environments and natural populations are essential for characterizing gene function. Therefore, *de novo* discovery of candidates in association tests would provide a valuable parallel strategy. One option is to identify candidate *regions* using populations with moderate levels of LD and high levels of genetic variation. *Populus* provides an excellent resource that fits these criteria: extensive zones of natural hybridization (Lexer *et al.*, 2004). All North American *Populus* species form hybrid swarms in zones where closely related but ecologically distinct species overlap (Eckenwalder, 1984). Some of these zones have been intensively characterized over many years, and it has become apparent that *Populus* plays a keystone role in these systems. Consequently, the genetic and phenotypic variation resulting from hybridization has consequences that are detectable at the ecosystem and even landscape scales (Whitham *et al.*, 2003). This raises the possibility of extending the frontier beyond molecular determinants of

adaptation of individual species to include the composition and functioning of communities and even ecosystems. Toward this end, the Department of Energy is now sequencing two of the major mycorrhizal symbionts of *Populus*, *Laccaria bicolor* and *Glomus intraradices* (Martin *et al.* 2004). The flood of data and discoveries will therefore continue as pioneer *Populus* researchers enter the age of ecosystem genomics.

Acknowledgements

Thanks to Steve Strauss, Malcolm Campbell, Francis Martin and Jerry Tuskan for their input on the manuscript and involvement in the meeting, to Holly Slater and Rich Norby for making it all happen, and to all of the discussion leaders, speakers and poster presenters for their outstanding contributions. Funding for the meeting was provided by the *New Phytologist* Trust and the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

Stephen P. DiFazio

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6422, USA
(tel +1 865 574 7844; fax +1 865 576 9939; email difazios@ornl.gov)

References

- Arimura G, Huber DPW, Bohlmann J. 2004. Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa* × *deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (-)-germacrene D synthase, PtdTPS1. *Plant Journal* 37: 603–616.
- Bhalerao R, Nilsson O, Sandberg G. 2003. Out of the woods: forest biotechnology enters the genomic era. *Current Opinion in Biotechnology* 14: 206–213.
- Braatne JH, Rood SB, Heilman PE. 1996. Life history, ecology, and reproduction of riparian cottonwoods in North America. In: Stettler, RF, Bradshaw, HD Jr, Heilman, PE, Hinckley, TM, eds. *Biology of Populus and its Implications for Management and Conservation*. Ottawa, Canada: NRC Research Press, 57–85.
- Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proceedings of the National Academy of Sciences* 101: 15255–15260.
- Brunner AM, Busov VB, Strauss SH. 2004. Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends in Plant Science* 9: 49–56.
- Brunner AM, Nilsson O. 2004. Revisiting tree maturation and floral initiation in the poplar functional genomics era. *New Phytologist* 164: 43–51.
- Eckenwalder JE. 1984. Natural intersectional hybridization between north American species of *Populus* (Salicaceae) in sections *Aigeiros* and *Tacamahaca*. II. Taxonomy. *Canadian Journal of Botany* 62: 325–335.
- Feder ME, Mitchell-Olds T. 2003. Evolutionary and ecological functional genomics. *Nature Reviews Genetics* 4: 651–657.
- Grunze N, Willmann M, Nehls U. 2004. The impact of ectomycorrhiza formation on monosaccharide transporter gene expression in poplar roots. *New Phytologist* 164: 147–155.
- Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler NC, Chen THH. 2003. From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany—Revue Canadienne de Botanique* 81: 1247–1266.
- Jackson RB, Linder CR, Lynch M, Purugganan M, Somerville S, Thayer SS. 2002. Linking molecular insight and ecological research. *Trends in Ecology and Evolution* 17: 409–414.
- Kirst M, Johnson AF, Baucom C, Ulrich E, Hubbard K, Staggs R, Paule C, Retzel E, Whetten R, Sederoff R. 2003. Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* 100: 7383–7388.
- Lexer C, Heinze B, Alia R, Rieseberg LH. 2004. Hybrid zones as a tool for identifying adaptive genetic variation in outbreeding forest trees: lessons from wild annual sunflowers (*Helianthus* spp.). *Forest Ecology and Management* 197: 49–64.
- Lynch M, Walsh B. 1998. *Genetics and Analysis of Quantitative Traits*. Sunderland, MA, USA: Sinauer Associates.
- Martin F, Tuskan GA, Difazio SP, Lammers P, Newcombe G, Podila GK. 2004. Symbiotic sequencing for the *Populus* mesocosm. *New Phytologist* 161: 330–335.
- Mitton JB. 1994. Molecular Approaches to Population Biology. *Annual Review of Ecology and Systematics* 25: 45–69.
- Neale DB, Savolainen O. 2004. Association genetics of complex traits in conifers. *Trends in Plant Science* 9: 325–330.
- Pilate G, Dejardin A, Laurans F, Leple JC. 2004. Tension wood as a model for functional genomics of wood formation. *New Phytologist* 164: 63–72.
- Remington DL, Ungerer MC, Purugganan M. 2001. Map-based cloning of quantitative trait loci: progress and prospects. *Genetical Research* 78: 213–218.
- Smith CM, Rodriguez-Buey M, Karlsson J, Campbell MM. 2004. The response of the poplar transcriptome to wounding and subsequent infection by a viral pathogen. *New Phytologist* 164: 123–136.
- Smouse PE, Sork VL. 2004. Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management* 197: 27–38.
- Sterky F, Bhalerao RR, Unneberg P, Segerman B, Nilsson P, Brunner AM, Charbonnel-Campaa L, Lindvall JJ, Tandre K, Strauss SH, Sundberg B, Gustafsson P, Uhlen M, Bhalerao RP, Nilsson O, Sandberg G, Karlsson J, Lundeberg J, Jansson S. 2004. A *Populus* EST resource for plant functional genomics. *Proceedings of the National Academy of Sciences, USA* 101: 13951–13956.
- Strauss SH, Martin FM. 2004. Poplar genomics comes of age. *New Phytologist* 164: 1–4.
- Tsai CJ, Hubscher SL. 2004. Cryopreservation in *Populus* functional genomics. *New Phytologist* 164: 73–81.
- Tuskan G, Difazio SP, Teichmann T. 2004. Poplar genomics is getting popular: The impact of the poplar genome project on tree research. *Plant Biology* 6: 2–4.
- Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, Shuster SM, Wimp GM, Fischer DG, Bailey JK, Lindroth RL, Woolbright S, Kuske CR. 2003. Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology* 84: 559–573.
- Yin TM, DiFazio SP, Gunter LE, Jawdy SS, Boerjan W, Tuskan GA. 2004. Genetic and physical mapping of *Melampsora* rust resistance genes in *Populus* and characterization of linkage disequilibrium and flanking genomic sequence. *New Phytologist* 164: 95–105.
- Zobel B, Talbert J. 1984. *Applied Forest Tree Improvement*. New York, USA: Wiley.

Key words: adaptive variation, complex trait, ecosystem genomics, genetic architecture, hybridization, model plant, poplar, population genetics.