



Poplar genome sequence: functional genomics in an ecologically dominant plant species

Amy M. Brunner, Victor B. Busov and Steven H. Strauss

Department of Forest Science, Oregon State University, Corvallis, OR 97331-5752, USA

In addition to their value for wood products, members of the genus *Populus* (poplars) provide a range of ecological services, including carbon sequestration, bioremediation, nutrient cycling, biofiltration and diverse habitats. They are also widely used model organisms for tree molecular biology and biotechnology. The sequencing of the poplar genome to an approximately 6× depth adds to a long list of important attributes for research. These include facile transformation, vegetative propagation, rapid growth, modest genome size and extensive expressed sequence tags. Here, we discuss how the genome sequence and transformability of poplar, together with its high levels of genetic and ecological diversity, are enabling new insights into the genetic programs controlling ontogeny, ecological adaptation and environmental physiology of trees.

The most prominent characteristics that make trees different from annual herbaceous plants are their self-supporting structure and long lifespans. Wood is both a supporting and a conductive structure, allowing trees to achieve immense sizes and to out-compete smaller plants for light, water and nutrients, enabling trees to live for centuries and even millennia. As a consequence, trees produce most of the terrestrial biomass and dominate most terrestrial ecosystems. Physiological characteristics that enable ecological dominance include: (i) extensive formation of secondary xylem; (ii) age-related phase changes in many aspects of morphology and physiology; (iii) mechanisms to cope with the highly variable biotic and abiotic stresses encountered during their long lifespan; (iv) the capacity for extremely long-distance transport of water, nutrients and macromolecules; and (v) systems for coordinating development and environmental responses over these long distances at the whole-plant level. The regulatory networks controlling tree development and environmental response are unlikely to be unique to trees; however, they have probably undergone substantial modification compared with annual and small-stature perennial plants. Understanding these subtleties at the molecular level will help in deciphering how trees survive and adapt, and how they can be bred more efficiently for diverse environmental and economic goals.

Our ability to understand the genes, regulatory networks and molecular mechanisms underlying development and adaptation in trees has been severely limited by their size and generation time, and our limited molecular genetic knowledge base. To make major inroads in this area, a model tree system is needed for which we have a complete genomics 'toolkit'. The genus *Populus* (i.e. poplars, including aspens and cottonwoods) has several attributes that have led to its emergence as the model system for tree molecular biology [1–3]. Most importantly, the power of poplar as a model plant has been dramatically enhanced by the recent sequencing of *Populus trichocarpa* (black cottonwood), the largest native angiosperm tree in western North America. Like most of the ~30 species in the genus, *P. trichocarpa* exhibits abundant genetic and adaptive variation throughout its natural range [4,5], which includes Alaska and Mexico. As with other dominant species, studies suggest that heritable genetic variation within *Populus* species has community and ecosystem consequences [6]. We discuss the natural attributes and genomic resources of poplars as tools for dissecting the molecular mechanisms underlying adaptation and ecophysiology of trees.

Poplar as a genomic system

The advantages of poplar as a genomic system for tree molecular biology have been reviewed extensively (Table 1) [1,2,7]. They are one of the few trees that can be transformed, regenerated, vegetatively propagated and grown to tree size (e.g. 3–10 m) in a sufficiently short time to enable their use in functional genomic studies (Figure 1). This is crucial if the comparative advantage of poplars over other plants as systems for tree functional genomics is to be realized. Traits such as crown structure, wood properties and phase change obviously benefit from early rapid growth to a substantial size.

Poplars were the first trees to be genetically transformed and regenerated; more studies have been published on transgenic poplars than on all other tree species combined [8]. Although unstable events can be identified [9], field tests have shown that transgene expression during tree growth and propagation is highly stable in most lines and that the level of somaclonal variation associated with gene transfer and vegetative propagation is modest (Figure 1) [10]. Large-scale functional genomics

Corresponding authors: Amy M. Brunner (amy.brunner@oregonstate.edu), Steven H. Strauss (steve.strauss@oregonstate.edu).

studies can therefore include extensive transformation events and vegetative replication to enable subtle physiological alterations to be identified. Although some genotypes are much easier to transform than others are, a wide range of species and hybrids have been successfully transformed [11]. This allows comparison of transgene-induced modifications among genotypes that are adapted to different environments.

Poplars are paleopolyploids (polyploids that have undergone diploidization) and all species have a haploid chromosome number of 19 [12]. Gene-family and restriction fragment length polymorphism mapping studies have identified many duplications of the poplar genome that appear to have occurred in the common ancestor of *Populus* or possibly the entire Salicaceae [13,14]. Nonetheless, they have a modest genome size (similar to rice and ~40 times smaller than pines), facilitating the isolation of genes (Table 1) [1]. The simplicity of its genome and its high ratio of genetic to physical map distance [1] will be increasingly important for gene identification as methods such as gene tagging and candidate gene association studies proceed.

Poplar genomic resources

A wide range of poplar genomic resources is publicly available and many of these can be directly assessed via GenBank and the worldwide web (Table 2). The International *Populus* Genome Consortium [based at Oak Ridge National Laboratory (ORNL) in the USA] is attempting to coordinate research internationally. Its website includes international science plans for poplar genome projects and useful links to diverse resources. The poplar nuclear and chloroplast genome sequence is available for downloading and searching at the Joint Genome Institute (JGI) and ORNL websites, respectively. Basal assembly of the draft shotgun-sequenced genome is proceeding with the aid of a bacterial artificial chromosome



Figure 1. Field trial of transgenic poplars (*Populus deltoides* × *Populus nigra*) in their second growing season that display low levels of somaclonal variation and rapid growth. The trees were grown under drip irrigation and fertilization in the north-western USA. Several ramets of nine independent transgenic lines and one non-transgenic control line of the same parent clone are planted randomly within alternating rows; all the trees in alternate rows contain only the non-transgenic line. Trees continued to show high uniformity in morphology for three growing seasons, after which the trial was terminated. (a) View between two adjacent rows (the man is 2 m in height). (b) Infrared aerial photograph of a section of the plantation.

fingerprint minimum-tiling path contributed by Genome British Columbia. The assembled genome contigs will be associated with large scaffolds representing the 19 *Populus* chromosomes using a series of framework MICROSATELLITE MARKERS (see Glossary) (<http://www.ornl.gov/ipgc/Links.htm>) distributed across the *Populus* chromosomes at 5 cM resolution that were developed at ORNL.

Similarly, basal annotation will be facilitated by several international expressed sequence tag (EST) collections. The Swedish *Populus* Genome Project was the first, and is the most advanced, integrated poplar genomics project in the world – it has already produced numerous transgenic

Table 1. Poplar genomic system

Attribute	Applications
Transformability	Many genotypes can be transformed and plants regenerated, some at high frequency; somaclonal variation minor; stable transgene expression predominant.
Vegetative propagation	Pedigreed and transgenic lines can be readily micropropagated <i>in vitro</i> , and often easily <i>in vivo</i> , with little maturation-associated somatic effects. The main <i>in vivo</i> methods are rooting of stem cuttings (cottonwoods) or 'root suckering' (aspens).
Fast growth rate	Tree phenotypes and wood can be produced in 1–3 years in greenhouse or field environments.
Rapid tree flowering	Flowering can be observed in the field after 3–6 years under rapid growth conditions, but can be induced within 1 year using transgenes or early flowering genotypes.
Ecological diversity	The ~30 species reside in diverse environments, from desert riparian to subalpine. Single species cover vast geographic and elevation ranges.
Intraspecific diversity	Levels of genetic diversity are high for molecular markers such as allozymes and simple sequence repeats, and for adaptive traits such as vegetative phenology.
Species hybridization	Many species with distinct ecological niches, including those from different taxonomic sections, can be hybridized and backcross or F ₂ generations produced and mapped using quantitative trait loci.
Genome size	Flow cytometry suggests a haploid genome size of 550 Mb [1]. Recent analyses of the genome sequence suggest 500 (± 20) Mb (D. Rokhsar, unpublished).
Polymorphic genetic markers	Several hundred highly polymorphic microsatellite markers have been developed, many of which have been mapped and function in diverse species.
Growing EST, genome, and microarray sequence databases	More than 200 000 ESTs have been identified from a wide range of tissues and laboratories, and used to produce a range of microarray chips for expression analysis (Table 2).
Phylogenomic proximity to well-studied angiosperms	Poplar is a eudicot within the rosid clade, facilitating comparative genomic inferences from studying <i>Arabidopsis</i> and the sequenced legumes such as <i>Medicago</i> and <i>Lotus</i> .

Glossary

Arbuscular mycorrhizal fungi: ubiquitous fungi of the order Glomales that form intracellular symbiosis with the roots of most land plants throughout the world; these fungi appear to have evolved with the land plants.

Ectomycorrhizal fungi: form intercellular symbioses with roots of woody plants that dominate forest ecosystems in boreal, temperate and Mediterranean regions.

Intrasectional hybrids: hybrids of two species within the same section of a genus (e.g. *Populus tremula* × *Populus tremuloides*).

Intersectional hybrids: hybrids of two species belonging to different sections of a genus (e.g. *Populus trichocarpa* × *Populus deltoides* is a hybrid of species from section *Tacamahaca* and section *Aigeiros* of the genus *Populus*).

Linkage disequilibrium: the nonrandom association of alleles at different loci within a population.

Micro-co-linearity: conservation between different species of the order of coding regions on >50 kb DNA fragments; microrearrangements such as insertions and deletions are often detected at this level of conservation.

Simple sequence repeats and microsatellite markers: loci where one or a few bases are tandemly repeated a variable number of times. SSRs show extensive length polymorphism and hence are widely used for DNA fingerprinting and diversity studies.

lines and microarrays [7,15]. Its large collection of ESTs has been clustered into a 25 000 unigene set, including 1000 full-length cDNA sequences, and it has also brought together a worldwide collection of over 200 000 ESTs (S. Jansson, pers. commun.). Genome British Columbia has produced 20 000 ESTs and plans to produce a total of 100 000 ESTs from normalized cDNA libraries and 5000 completely sequenced cDNAs from full-length enriched cDNA libraries (J. Bohlmann, C. Douglas, B. Ellis and K. Ritland, pers. commun.). To complete the basal annotation, researchers at JGI, ORNL and the University of Ghent will use three autonomous gene-calling algorithms, and the assembled and annotated genome with physical and genetic map interfaces should be accessible through a website curated by JGI in early 2004 (G. Tuskan, pers. commun.)

Two conspicuous voids in poplar genomic resources are well characterized, publicly available genotype collections and a real or virtual stock center to maintain and direct distribution of materials. Because of the high heterozygosity of poplars, the limited lifespan of their small seeds, their vegetative longevity and their amenability to vegetative propagation, vegetative rather than seed materials are likely to be preferred. For valuable collections, this is likely to include cryopreservation. With the diversity of environments, materials and traits of interest, as well as quarantine and regulatory restrictions to shipment, a virtual stock center in which materials are housed in several places but tracked and coordinated in one place would seem to be logical. Given the importance of transgenics for functional genomics, this center might logically be linked to a transformation service center.

Comparative genomics

Poplars belong to the rosids and so are well positioned for comparative analysis with other intensively studied and sequenced genomes (Figure 2a). Phylogenetic analysis of gene families can identify putative orthologs to genes characterized in *Arabidopsis* or other annuals, providing clues to gene function, and can reveal lineage-specific duplications (Figure 2b). Comparative mapping has revealed extensive genome co-linearity between species in the same family and microsyntenic relationships between distantly related species [16]. Regions of MICRO-CO-LINEARITY have already been identified between poplar and *Arabidopsis* [17]. Syntenic relationships can provide unambiguous identification of orthologs and extend genetic maps established in one species to related species. Large EST databases are being developed for an increasing number of trees (Figure 2a). With its complete

Table 2. Poplar genomic resources^a

Institution	Worldwide-web URL	Key resources
International <i>Populus</i> Genome Consortium, Oak Ridge National Laboratory, USA	http://www.ornl.gov/ipgc/	International research coordination, simple sequence repeat databases, research protocols, links
USA Department of Energy, Joint Genome Institute, <i>Populus trichocarpa</i>	http://genome.jgi-psf.org/poplar0/poplar0.home.html	Genome sequence, downloads, Blast, basic genome annotation
Swedish <i>Populus</i> Genome Project	http://popel.fysbot.umu.se/	EST databases, microarrays, aspen wood formation and physiology
PoplarDB/Lignome, INRA, France	http://mycor.nancy.inra.fr/poplar0/	ESTs of roots, wood, leaves; mycorrhizal interactions
Oak Ridge National Laboratory Computational Biology, USA	http://genome.ornl.gov/poplar_chloroplast/index.html	Annotated <i>Populus</i> chloroplast genome sequence
Genome Canada, University of British Columbia and Laval University	http://www.bcgsc.ca/gc/poplar; http://www.arborea.ulaval.ca/en/	Physical maps, BAC-end sequences, ESTs, microarrays, simple sequence repeats, transgenics, activation tagged mutants
Dendrome, US Forest Service and University of California, Davis	http://dendrome.ucdavis.edu/	Diverse tree genome databases and links, scientist directories
Popyomics, Fifth Framework Programme of the EU	http://www.soton.ac.uk/~popyomic/	Pedigrees for mapping yield and disease resistance
Michigan Technological University Plant Biotechnology Research Center	http://forest.mtu.edu/research/pbrc/	EST databases, cellulose and lignin
Oregon State University Tree Genetic Engineering Research Cooperative	http://www.fsl.orst.edu/tgerc/	Transformation services center, activation tagged mutants, floral genes
Tree Biotechnology 2003, International Meeting in June 2003, Sweden	http://www.treebiotech2003.norrnod.se/	Abstracts of recent research in poplar genomics and biotechnology

^aURLs correct on 11 November 2003.

genomics toolkit, poplar will be a valuable reference species to facilitate the study of other tree species.

Whole genome sequences enable computational approaches to identifying *cis* regulatory elements based on phylogenetic conservation and similarity in expression. Poplar genomic sequences have already been used to identify regulatory motifs in the floral development gene *AGAMOUS* from various eudicots [18]. Comparisons of complete genome sequences combined with various molecular genetic approaches are particularly powerful for advancing comparative biology. For example, the complete genome sequences of rice and *Arabidopsis* have, together with genetic studies, enabled comparisons of the regulatory networks controlling flowering between these distantly related angiosperms [19].

The addition of a dominant tree species to the pantheon of complete genome sequences promises many new insights from comparisons of signaling networks, such as those involved in symbiotic plant-root–microorganism interactions. Studies of intracellular symbioses indicate that signaling between nitrogen-fixing soil bacteria and legumes might have evolved from pre-existing pathways that regulate the more ancient and widespread symbioses between ARBUSCULAR MYCORRHIZAL FUNGI and most plant

roots [20]. In contrast to arbuscular mycorrhizal fungi, ECTOMYCORRHIZAL FUNGI form intercellular symbioses with a more limited number of plants, but many of these species are ecologically dominant woody angiosperm and gymnosperm species, including poplars [21]. Poplar functional genomics will enable direct insights into the fundamental signaling pathways involved in ectomycorrhizal symbiosis and is also likely to provide insights into the mechanisms and evolution of all plant-root–microbe symbioses [21,22].

Functional genomics

Apart from the use of genomics to aid wood biotechnology research, which has been well reviewed elsewhere [7], the primary motivation for studying poplars is to understand how forest and tree diversity, adaptation to environment, ecological services and productivity are controlled. Fundamental cellular and physiological aspects of gene function in plants can be far more precisely analyzed in *Arabidopsis*, rice and other annual model plants. Thus, a key challenge for poplar research is how to conduct affordable experiments that enable gene-level inferences to be made about tree-level attributes such as wood function, vegetative and reproductive phase change and

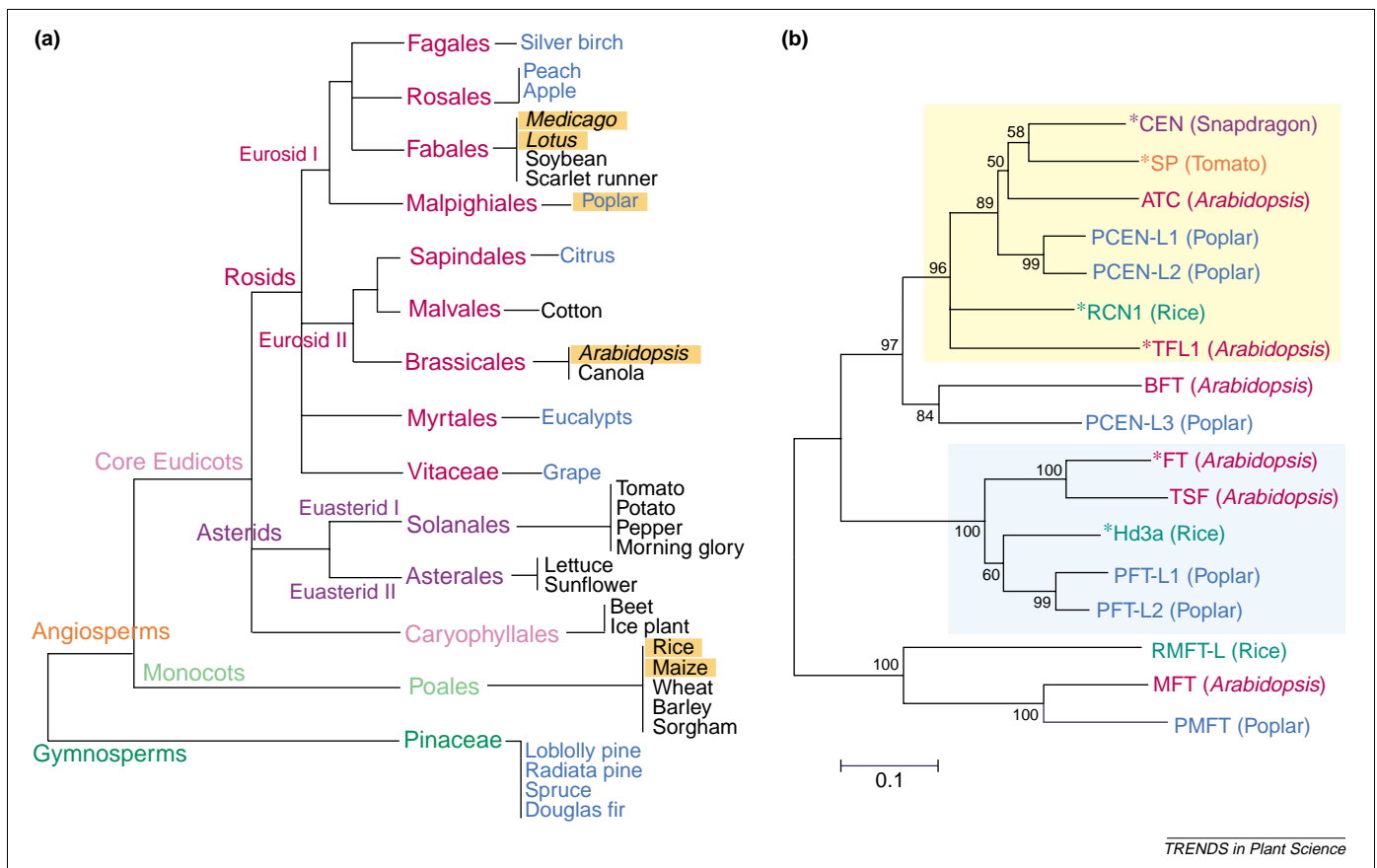


Figure 2. Phylogenetic relationships of poplars and of a poplar gene family, showing the power of comparative genomics approaches in poplar. (a) Angiosperm phylogeny based on Soltis *et al.* [59], showing only orders that include genera with > 10 000 expressed sequence tags (ESTs) in dbEST. Also included are trees for which > 10 000 ESTs are in the process of being generated or for which there are large, currently proprietary, EST databases (i.e. eucalypts and radiata pine). Woody plants are in blue type and species for which complete genome sequences are, or soon will be, available are in orange boxes. (b) Phylogenetic tree of CENTRORADIALIS (CEN)/TERMINAL FLOWER1 (TFL1)-like regulatory proteins [60] produced by the neighbor-joining method. Bootstrap values (percentages based on 1000 replicates) are shown at nodes. All six *Arabidopsis* family members, *Populus trichocarpa* members present in GenBank or identified in the genome sequence, and selected genes from other species are included. The subclade highlighted in yellow includes genes that have been shown to maintain the indeterminate inflorescence or vegetative phases, whereas the subclade highlighted in pale blue includes genes that promote the transition to flowering.

phenology, abiotic stress tolerance, and root and shoot biotic interactions.

Although some types of genomic studies can be conducted in the greenhouse, they are costly, limit population and tree sizes, and the trees have abnormal physiology compared with those grown in outdoor environments. Because of the intensive government regulation of transgenic organisms in many countries and risks of vandalism (even where non-flowering transgenic trees are used), there can be difficult legal and political issues to face in conducting field trials. In spite of some pioneering field studies on lignin modification [23], these difficulties have limited field trials in many parts of Europe; most transgenic studies are therefore taking place in the USA, China and a few other countries (<http://www.isb.vt.edu/>).

Because of the long generation time, difficulty of inbreeding and rarity of mendelian traits in wild populations, there are no well-characterized mutant collections of forest trees. The generation time and high genetic load make inbreeding to reveal natural recessive mutants, or those induced by mutagens or transgenes, experimentally difficult. Transformation, by enabling dominant, directed changes in expression of single genes, thus provides a revolutionary increase in the ability to link single genes to phenotypes in trees [24].

'Reverse genetics', in which a gene sequence is first obtained and then its function learned via directed alteration in transgenics, is the route taken in most transgenic experiments with trees. Typically, gene expression is either suppressed via various forms of RNA-mediated silencing [25], or gene expression is elevated via use of a strong promoter or enhancer. This method allows genes to be carefully selected for study based on knowledge of gene function in other organisms, requiring that far fewer mutants be produced than in random forward genetics mutagenesis programs. Nonetheless, forward genetics approaches such as activation tagging and enhancer traps – because they allow dominant tagged mutations to be produced and are not constrained by knowledge of gene function from herbaceous plants – are also starting to be used in poplars [26].

High throughput methods for studying tree and wood phenotypes

To study the arboreal characteristics of poplars, large plants and preferably trees grown in outdoor environments are desirable. This poses substantial logistical problems for collecting phenotype information. Trees, particularly fast-growing poplars, take up a great deal of space and must be grown for several years without competition between genotypes confounding results. For example, to analyze a plantation of ~ 10 000 independent transgenic events replicated four times would require ~ 32 hectares. Moreover, some important traits, such as vegetative phenology and reactions to episodic climate or biotic events, requires that many thousands of dispersed phenotypes be determined over as short a time period as one week – a near-impossible task for observers on the ground. Other traits, such as crown structure, are

spatially or chemically complex and thus difficult to measure with standard physiochemical techniques. Assessments of these traits in large plantations are therefore likely to require remote-sensing technology, preferably including the multispectral sensors used for ecophysiological studies of canopies [27]. Finally, wood is chemically and structurally complex and difficult to sample. Fortunately, high-throughput methods for study (such as the use of infrared spectral analysis and mass spectrometry for wood chemistry and microdensitometry, image analysis and X-ray diffraction for wood anatomical properties) enable rapid screens of wood for genetic alterations [28,29]. The CSIRO Forestry and Forest Products Program (Australia) offers a variety of high-throughput wood structure analyses through its SilviScan[®] service (<http://www.fpp.csiro.au/tigr/silviscan/>).

Poplar as a model system to study adaptation

Natural variation provides useful tools for dissecting physiological mechanisms and poplars exhibit abundant variation in virtually every adaptive trait examined. Poplars are widely distributed across the Northern Hemisphere from the tropics to beyond the Arctic Circle, and these ranges span highly contrasting environments [4]. For example, trembling aspen (*Populus tremuloides*) is the most geographically widespread native tree species in North America, occurring from Mexico to north-western Alaska and from the Atlantic to the Pacific Ocean. Partly as a consequence of their broad ecological ranges, poplars are highly polymorphic for photoperiodic responses [30], crown architecture [31], cold hardiness [32] and wood structure [33].

Like other tree species with continuous geographic distributions, poplars display high levels of genetic diversity within populations at molecular marker loci [34,35]. A random sample of 15 SIMPLE SEQUENCE REPEAT loci in three wild populations (mean $n = 200$ trees) of *P. trichocarpa* in Oregon showed an average number of alleles per locus of 20 and an observed heterozygosity of 77% (S. DiFazio and S. Strauss, unpublished). The single nucleotide polymorphism rate in the genome sequence (which was produced from a wild, highly heterozygous tree) is ~0.5%. This polymorphism rate is greater than that of the human genome by about a factor of five but is comparable to the sequenced genome of the Japanese pufferfish (*Fugu rubripes*) (D. Rokhsar, unpublished). Because of the high rates of outcrossing in wind-pollinated dioecious poplars, most genetic diversity is allocated within populations. The high genetic polymorphism and limited population differentiation facilitates association mapping of genetic and phenotypic variation.

Molecular ecogenetics

Quantitative trait locus (QTL) approaches have been widely used to dissect adaptive responses in poplars [1,2,36]. The availability of the poplar genome sequence has accelerated the development of high-density genetic markers and will thus aid in making physical linkages between QTLs and genome regions. However, the requirement for large mapping pedigrees and qualitative traits to define genomic regions that contain a limited number of genes make this

approach difficult to implement for many of the adaptive traits important to trees.

An alternative approach, association mapping, takes advantage of the high degree of recombination in natural populations that has occurred over hundreds of generations [37]. The success of this approach depends on the amount of LINKAGE DISEQUILIBRIUM, population structure and other factors [38]. However, the low linkage disequilibrium expected in poplars might require that many thousands of markers are used for a full genome scan in a natural population. Candidate gene association mapping (in which specific genes of interest – rather than the entire genome – are scanned for association with traits) circumvents these problems but limits studies to known gene-physiological hypotheses [39].

By using a range of annotation and comparative informatics approaches, the genome sequence will ultimately reveal most of the useful genetic markers and genes, and their positions in the genome. High-throughput single nucleotide polymorphism and microsatellite detection and scoring [40,41] should provide new options for large-scale studies of genomic diversity. In addition, microarray methods provide new ways of studying gene expression in relation to development and adaptation on a genomic scale. For example, a recent study in mammals and maize identified QTLs associated with variation in RNA expression [42].

Adaptive traits important to woody plants

Woody plants from the boreal, temperate and subtropical regions have evolved mechanisms to alternate periods of active shoot growth and vegetative quiescence that closely track local microclimates. Although dormancy – formally defined as a temporary suspension of the visible growth of plant structures that contain meristems – also occurs in annual plants, it is a more complex developmental process in perennial species [43].

In temperate zone trees, dormancy is a multi-stage cycle that occurs over several months. A key element is the formation of a dormant bud that can endure extreme freezing and dehydration stress during winter conditions and that enables rapid growth in spring following an extended chilling period. The molecular mechanisms underlying this sequence of events is still poorly understood but there is already a considerable body of molecular genetic research in poplar that demonstrates its value as an experimental system for fundamental study. Jarmo Schrader *et al.* [44] used poplar to demonstrate a role for polar auxin transport genes during vascular cambium development and entry into dormancy. Candidate genes have been shown to localize with QTLs for poplar bud phenology [36]. Study of transgenic poplars revealed roles for *ABSCISIC ACID-INSENSITIVE3* in bud set [45] and for *PHYTOCHROME A* in short-day-induced growth cessation and cold acclimation [46].

From ecological and evolutionary perspectives, wood formation can be viewed as an adaptive mechanism that enables trees to secure a dominant position in ecosystems, promoting their survival, reproduction and dispersal of propagules. Wood has many functions, including water transport, mechanical support, regeneration after injury

and storage of organic compounds, water and gases [47]. The basis for these functions is a complex, environmentally responsive developmental program that produces a structure composed of cells oriented both axially (e.g. xylem vessels) and radially (e.g. rays) (Figure 3) [48].

There is extensive physiological and genetic variation in wood structure, physical properties and chemical composition [33,49]. However, its importance for adaptation is poorly understood. For example, the radius of xylem cells is important for water conductivity and consequently to drought tolerance. Larger tracheid lumens allow higher conductivity but greatly increase susceptibility to embolism [50]. Self-supported trees and shrubs have lower water conductivity but wood with higher mechanical strength than externally supported vines do [51,52]. This suggests that there is a powerful trade-off between mechanical strength, rate of water conductivity and maximum rate of growth that must be respected when large changes in wood properties are contemplated during molecular domestication. The identification of new genes involved in wood formation [7,15] should provide many new tools to analyze the adaptive significance of variation in wood structure.

Epigenetic regulation and the life history of trees

The fundamental importance of epigenetic regulation to the life history of trees has long been inferred, but has largely been intractable to study. The poplar genomics toolkit now offers the opportunity to change this situation radically. Functional characterization of all the *Arabidopsis* and most of the maize chromatin regulatory genes is in progress (<http://www.chromdb.org/>). The results of this project will inform studies of how epigenetic regulation controls tree development.

Many studies have shown that chromatin regulators, which mediate epigenetic control of transcription, are key regulators of plant development [53]. Epigenetic

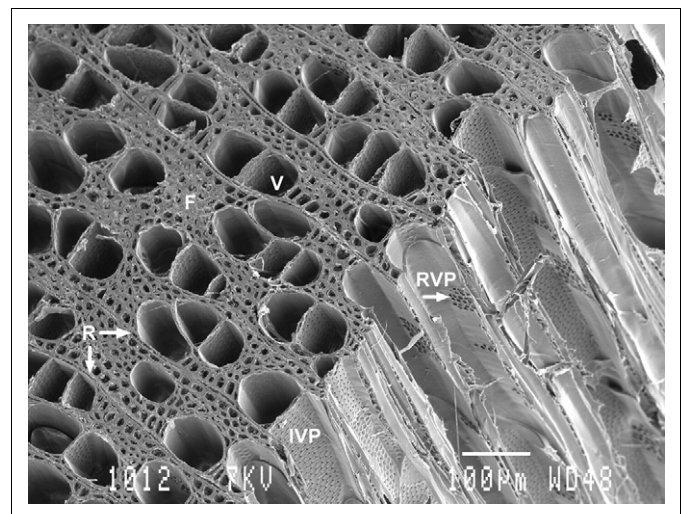


Figure 3. Scanning electron micrograph of wood from *Populus tremuloides* showing two section planes. Top left: the transverse or cross-section plane (cut horizontally across vascular elements). Bottom right: the tangential plane (cut vertically in parallel with the vascular elements and vascular cambium). Abbreviations: F, fibers; IVP, inter-vessel pits; R, rays; RVP, ray-vessel pits; V, vessels. Micrograph courtesy of James Drummond and Simon Potter (Pulp and Paper Research Institute of Canada, Vancouver, BC, Canada).

regulation is also crucial for genome maintenance through activities such as silencing of retrotransposons. The longevity and size of trees allows greater opportunity for transposon activation and epiallele formation caused by stresses or other factors [54]. However, the extent to which epialleles occur and are a significant source of adaptive variation in trees are unknown.

In contrast to annuals, trees maintain developmental phases for many traits for years, including wood characteristics, pest resistance, competence for flowering and crown architecture [55]. In addition to multi-year phases, trees also exhibit recurring seasonal phase changes, such as the transition from early wood to late wood. They also maintain stable within-tree developmental gradients, such as greater juvenility towards the base compared with the tips of the crown. Although some changes are clearly part of a developmental program (e.g. the onset of flowering), the causes of other changes that occur with age, particularly those associated with the old-growth stage of trees, are uncertain and could be because of size or cumulative environmental stresses. The consequences of age- and size-related alterations in developmental patterns for tree fitness are expected to be great; trees face different challenges to survival and reproduction at different life stages [56].

The amenability of poplars to epigenetic studies is enhanced by several natural attributes. There is rapid somatic reversibility in developmental phases in several species and this capacity varies between traits, genotypes and species. For example, mature branch segments of several cottonwoods, including the sequenced *P. trichocarpa*, root easily and then rapidly revert to juvenile characteristics (e.g. the first newly formed leaves are juvenile). Epigenetic changes have been shown to occur following polyploid formation and hybridization, and might contribute to heterosis [57]. Both INTRASECTIONAL HYBRIDS and INTERSECTIONAL HYBRIDS of poplars are widespread in nature, can include triploids and are widely planted in forestry, farm and urban environments. Moreover, studies of hybrid poplar populations indicate that they act as evolutionary filters, preventing introgression of most genes but allowing others to introgress throughout the recipient species range [58].

The tree-lined road ahead

Poplars are well suited to be a 'bridge species' for molecular ecophysiology – a vehicle with which the growing knowledge of *Arabidopsis* and other model annual plants can be extended to help understand the function of trees and forests. Although poplars have many advantages as model organisms, continued development is needed if their potential is to be realized. Key needs include: a community-wide genome chip to standardize RNA expression analyses and reporting of data; a transformation center where healthy transgenic plants can be produced to order at low cost; virtual and physical stock centers to store and distribute valuable genotypes; standardization of high throughput, low-cost methods for wood and metabolic characterization; and a system for rapid cycling and production of sexual progeny. Recent poplar transgenic studies have identified genes that

appear to be capable of inducing functional flowers on juvenile trees (O. Nilsson, pers. commun.). Finally, a great deal can be done to develop the transformability of poplar into a high throughput, efficient system, including improvements to RNA interference 'knockout' systems, the development of methods for controlled transgene expression and excision, and the development of reliably sterile backgrounds that enable long-term trials of transgenic mutants. None of these appear to be insurmountable obstacles if there is global collaboration in deciphering how the functioning of genomes can give rise to healthy and productive forests.

References

- Bradshaw, H.D. *et al.* (2000) Emerging model systems in plant biology: poplar (*Populus*) as a model forest tree. *J. Plant Growth Regul.* 19, 306–313
- Taylor, G. (2002) *Populus: Arabidopsis* for forestry. Do we need a model tree? *Ann. Bot.* 90, 681–689
- Wullschlegel, S.D. *et al.* (2002) Genomics and the tree physiologist. *Tree Physiol.* 22, 1273–1276
- Dickmann, D.I. (2001) An overview of the genus *Populus*. In *Poplar Culture in North America (Part 2)* (Dickmann, D.I. *et al.*, eds), pp. 1–42, NRC Research Press
- Eckenwalder, J.E. (1996) Systematics and evolution of *Populus*. In *Biology of Populus and its Implications for Management and Conservation* (Stettler, R.F. *et al.*, eds), pp. 7–32, NRC Research Press
- Witham, T.G. *et al.* (2003) Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology* 84, 559–573
- Bhalerao, R. *et al.* (2003) Out of the woods: forest biotechnology enters the genomic era. *Curr. Opin. Biotechnol.* 14, 206–213
- Peña, L. and Seguin, A. (2001) Recent advances in the genetic transformation of trees. *Trends Biotechnol.* 19, 500–506
- Kumar, S. and Fladung, M. (2001) Gene stability in transgenic aspen (*Populus*). II. Molecular characterization of variable expression of transgene in wild and hybrid aspen. *Planta* 213, 731–740
- Meilan, R. *et al.* (2002) Stability of herbicide resistance and *GUS* expression in transgenic hybrid poplars (*Populus* sp.) during several years of field trials and vegetative propagation. *HortScience* 37, 1–4
- Han, K.H. *et al.* (2000) An *Agrobacterium* transformation protocol effective in a variety of cottonwood hybrids (genus *Populus*). *Plant Cell Rep.* 19, 315–320
- Soltis, P.S. and Soltis, D.E. (2000) The role of genetic and genomic attributes in the success of polyploids. *Proc. Natl. Acad. Sci. U. S. A.* 97, 7051–7057
- Brunner, A.M. *et al.* (2000) Structure and expression of duplicate *AGAMOUS* orthologs in poplar. *Plant Mol. Biol.* 44, 619–634
- Bradshaw, H.D. *et al.* (1994) Molecular genetics of growth and development in *Populus*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theor. Appl. Genet.* 89, 167–178
- Hertzberg, M. *et al.* (2001) A transcriptional roadmap to wood formation. *Proc. Natl. Acad. Sci. U. S. A.* 98, 14732–14737
- Schmidt, R. (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Mol. Biol.* 48, 21–37
- Stirling, B. *et al.* Comparative sequence analysis between orthologous regions of the *Arabidopsis* and *Populus* (poplar) genomes reveals substantial synteny and microcollinearity. *Can. J. For. Res.* (in press)
- Hong, R.L. *et al.* (2003) Regulatory elements of the floral homeotic gene *AGAMOUS* identified by phylogenetic footprinting and shadowing. *Plant Cell* 15, 1296–1309
- Izawa, T. *et al.* (2003) Comparative biology comes into bloom: genomic and genetic comparisons of flowering pathways in rice and *Arabidopsis*. *Curr. Opin. Plant Biol.* 6, 113–120
- Kistner, C. and Parniske, M. (2002) Evolution of signal transduction in intracellular symbiosis. *Trends Plant Sci.* 7, 511–518
- Martin, F. *et al.* (2001) Frontiers in molecular mycorrhizal research – genes, loci, dots and spins. *New Phytol.* 150, 499–507
- Kohler, A. *et al.* (2003) The poplar root transcriptome: analysis of 7,000 expressed sequence tags. *FEBS Lett.* 542, 37–41

- 23 Pilate, G. *et al.* (2002) Field and pulping performance of transgenic trees with altered lignification. *Nat. Biotechnol.* 20, 607–612
- 24 Bradshaw, H.D. Jr and Strauss, S.H. (2001) Breeding strategies for the 21st century: domestication of poplar. In *Poplar Culture in North America (Part 2)* (Dickmann, D.I. *et al.*, eds), pp. 383–394, NRC Research Press
- 25 Waterhouse, P.M. and Helliwell, C.A. (2003) Exploring plant genomes by RNA-induced gene silencing. *Nat. Rev. Genet.* 4, 29–38
- 26 Busov, V.B. *et al.* (2003) Activation tagging of a dominant gibberellin catabolism gene (*GA 2-oxidase*) from poplar that regulates tree stature. *Plant Physiol.* 132, 1283–1291
- 27 Yoder, B.J. and Pettigrew-Crosby, R.E. (1995) Predicting nitrogen and chlorophyll from reflectance spectra (400–2500 nm) at leaf and canopy scales. *Remote Sensing Environ.* 53, 199–211
- 28 Kelly, S. *et al.* (2002) Use of NIR and MBMS coupled with multivariate analysis for detecting the chemical changes associated with brown-rot biodegradation of spruce wood. *FEMS Microbiol. Lett.* 209, 107–111
- 29 Tuskan, G.A. *et al.* (1999) Two high-throughput techniques for determining wood properties as part of a molecular genetics analysis of loblolly pine and hybrid poplar. *Appl. Biochem. Biotechnol.* 77–79, 55–65
- 30 Howe, G.T. *et al.* (1995) Photoperiodic responses of a northern and southern ecotype of black cottonwood. *Physiol. Plant.* 93, 698–708
- 31 Dunlap, J.M. *et al.* (1995) Genetic variation and productivity of *Populus trichocarpa* and its hybrids. VIII. Leaf and crown morphology of native *P. trichocarpa* clones from four river valleys in Washington. *Can. J. Forest Res.* 25, 1710–1724
- 32 McCamant, T. and Black, R.A. (2000) Cold hardiness in coastal, montane, and inland populations of *Populus trichocarpa*. *Can. J. Forest Res.* 30, 91–99
- 33 Yu, Q. *et al.* (2001) Genetic control of wood physicochemical properties, growth, and phenology in hybrid aspen clones. *Can. J. Forest Res.* 31, 1348–1356
- 34 Stevens, T.M. *et al.* (1999) Genetic variation in postfire aspen seedling in Yellowstone National Park. *Mol. Ecol.* 8, 1769–1780
- 35 Yeh, F.C. *et al.* (1995) RAPD variation within and among natural populations of trembling aspen (*Populus tremuloides* Michx.) from Alberta. *J. Hered.* 86, 454–460
- 36 Frewen, B.E. (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics* 154, 837–845
- 37 Remington, D.L. *et al.* (2001) Structure of linkage disequilibrium and phenotypic associations in maize genome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11479–11484
- 38 Rafalski, A. (2002) Application of single nucleotide polymorphism in crop genetics. *Curr. Opin. Plant Biol.* 5, 94–100
- 39 Garcia-Gil, M.R. *et al.* (2003) Nucleotide diversity at two polymorphic loci along a latitudinal cline in *Pinus sylvestris*. *Mol. Ecol.* 12, 1195–1206
- 40 Shi, M.M. (2001) Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin. Chem.* 47, 164–172
- 41 Borevitz, J.O. *et al.* (2003) Large-scale identification of single-feature polymorphisms in complex genomes. *Genome Res.* 13, 513–523
- 42 Schadt, E.E. *et al.* (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302
- 43 Rohde, A. *et al.* (2000) Molecular aspects of bud dormancy in trees. In *Molecular Biology of Woody Plants* (Vol. 1) (Jain, S.M. and Minocha, S.C., eds), pp. 89–134, Kluwer Academic Publishers
- 44 Schrader, J. *et al.* (2003) Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10096–10101
- 45 Rohde, A. *et al.* (2002) *PtABI3* impinges on the growth and differentiation of embryonic leaves during bud set in poplar. *Plant Cell* 14, 1885–1901
- 46 Welling, A. *et al.* (2002) Independent activation of cold acclimation by low temperature and short photoperiod in hybrid aspen. *Plant Physiol.* 129, 1633–1641
- 47 Gartner, B.L. (1995) *Plant Stems: Physiology and Functional Morphology*, Academic Press
- 48 Mellerowicz, E. *et al.* (2001) Unravelling cell wall formation in the woody dicot stem. *Plant Mol. Biol.* 47, 239–274
- 49 Atwood, R.A. *et al.* (2002) Genetic parameters and gains for growth and wood properties in Florida source loblolly pine in the southeastern United States. *Can. J. Forest Res.* 32, 1025–1038
- 50 Lo Gullo, M.A. and Salleo, S. (2003) Different vulnerabilities of *Quercus ilex* L. to freeze- and summer-induced xylem embolism: an ecological interpretation. *Plant Cell Environ.* 16, 511–519
- 51 Ewers, F.W. (1985) Xylem structure and water conduction in conifer trees, dicot trees, and lianas. *IAWA Bull.* 6, 309–317
- 52 Gartner, B.L. (1991) Stem hydraulic properties of vines vs. shrubs of western poison oak, *Toxicodendron diversilobum*. *Oecologia* 87, 180–189
- 53 Wagner, D. (2003) Chromatin regulation of plant development. *Curr. Opin. Plant Biol.* 6, 20–28
- 54 Finnegan, E.J. (2002) Epialleles: a source of random variation in times of stress. *Curr. Opin. Plant Biol.* 5, 101–106
- 55 Brunner, A.M. *et al.* (2003) Controlling maturation and flowering for forest tree domestication. In *Transgenic Plants: Current Innovations and Future Trends* (Stewart, C.N., ed.), pp. 9–44, Horizon Scientific Press
- 56 Day, M.E. *et al.* (2002) Age- and size-related trends in woody plant shoot development: regulatory pathways and evidence for genetic control. *Tree Physiol.* 22, 507–513
- 57 Osborn, T.C. *et al.* (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* 19, 141–147
- 58 Martinsen, G.D. *et al.* (2001) Hybrid populations selectively filter gene introgression between species. *Evolution Int. J. Org. Evolution* 55, 1325–1335
- 59 Soltis, D.E. *et al.* (2000) Angiosperm phylogeny inferred from 18S rDNA, *rbcL* and *atpB* sequences. *Bot. J. Linn. Soc.* 133, 381–461
- 60 Kobayashi, Y. *et al.* (1999) A pair of related genes with antagonistic roles in mediating flowering signals. *Science* 286, 1960–1962

News & Features on BioMedNet

Start your day with *BioMedNet's* own daily science news, features, research update articles and special reports. Every two weeks, enjoy *BioMedNet Magazine*, which contains free articles from *Trends*, *Current Opinion*, *Cell* and *Current Biology*. Plus, subscribe to Conference Reporter to get daily reports direct from major life science meetings.

<http://news.bmn.com>