FORESTRY IDEAS, 2013, vol. 19, No 2 (46): 107–128

THE PROMISE OF GENOMICS AND BEYOND

Victor Busov

School of Forest Resources and Environmental Sciences, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, USA. Email: vbusov@mtu.edu

Received: 24 July 2013 Accepted: 24 September 2013

Abstract

This review aims at providing a short synopsis of the history, advances and future of genomics sciences in various fields of research with focus on areas pertaining to forest tree species and ecosystems. It also provides a broader outlook of the place of genomics and other 'omics' approaches like transcriptomics, proteomics and metabolomics in future biology research at various levels of organization, particularly on system level knowledge at organismal and ecosystem scales. Specifically, emphasis is placed on new sequencing technologies that open the gate for thorough exploration of the DNA and RNA world from a single cell and individual to population and ecosystem scales. These and other technological innovations are holding great promise for advanced breeding, deeper knowledge of developmental and physiological mechanisms and better understanding of the processes occurring at ecosystem levels.

Key words: Systems biology, genetics, functional genomics, next-generation sequencing.

Brief History and Definitions

Genomics is a field of science dealing with a large number, if not all, genes and, more recently, the DNA sequence of the entire genome. This is in sharp contrast with preceding endeavors focusing typically on one or a few genes at a time. The dawn of genomics dates back to the late 1980s. Yet, the idea of sequencing and understanding the genetic information in its entirety is not new and is probably as old as the realization that DNA is the carrier of the hereditary information. However, it was only in the mid to late 1980s when the convergence of several major technological advances provided the foundations for the modern genomics field. These included recombinant DNA (Jackson et al. 1972), restriction enzymes (Luria and Human 1952), reverse transcriptase, polymerase chain reaction (Saiki et al. 1985) and Sanger sequencing (Sanger et al. 1977). The latter, followed by the automation of the method (Martingallardo et al. 1992), was the catalyst of the ensuing explosion of genomics sequencing data, resulting in complete sequences of initially prokaryotic and simple eukaryotic genomes (Blattner et al. 1997, Fleischmann et al. 1995, Goffeau et al. 1996), then closely followed by several more complex eukaryotic genomes like *Caenorhabditis elegans* (Maupas, 1900, *C. elegans* Sequencing Consortium 1998), *Drosophila melanogaster* (Meigen, 1830, Adams et al. 2000), *Arabidopsis thaliana* (L.) Heynh. (Kaul et al. 2000), and culminating in the publication of the human genome at the turn of the millennium (Lander et al.

2001, Venter et al. 2001). Particularly instrumental in the sequencing of these genomes and gene discovery in organisms with more complex and large genomes is the development and application of the Expressed Sequence Tag (EST) technology in the early 1990s (Adams et al. 1991). ESTs represent random partial sequences of small stretches of complementary DNA (cDNA) (derived from mRNA through reverse transcription) that are typically captured in plasmid vectors constituting what is known as cDNA libraries. The EST technology has been the workhorse of gene discovery in a number of species, including forest trees, for many years (Allona et al. 1998, Hertzberg et al. 2001, Kirst et al. 2003). Because of the complexity of conifer genomes, even for current sequencing capacities, ESTs have allowed gathering of useful genomics data by reducing the complexity and size of the sequencing effort that would have been impossible to afford by a whole-genome approach. ESTs have also proved extremely useful for 'training' the models for gene prediction, following whole-genome sequencing. Early realizations that the *de novo* gene prediction (based purely on sequence inferences) was extremely error-prone made the ESTs an invaluable resource to validate and 'educate' gene prediction algorithms (Martingallardo et al. 1992). Lessons learned from assembly of ESTs also led to the computational approaches used for assembly of whole-genome sequencing using the shotgun approach that is now accepted as a standard for whole-genome sequencing.

The know-how learned during the sequencing of the human genome opened the floodgates for sequencing of an array of other eukaryotic genomes from every nook and cranny of life. Plant genomes were among the first to be sequenced starting with the model *Arabidopsis thaliana* (Kaul et al. 2000), and even preceding the release of the human genome. This was closely followed by the publication of the rice genome (Goff et al. 2002) and ultimately the first tree and third plant genome: this of a poplar tree (*Populus trichocarpa* Torr. & A.Gray – Nisqually1 clone from the Pacific Northwest of the United States) (Tuskan et al. 2006).

Breaking the Confines of Model Organisms

Despite the significant advances in sequencing and assembly of genomes, these ventures remained confined to single genotypes and species of significance as model organisms or of significant economic, environmental and epidemiological importance. It is not coincidental that the first three sequenced eukaryotic genomes were of such organisms – *Saccharomyces cerevisiae* (Meyen & Hansen, 1883) (yeast), *Caenorhabditis elegans* (nematode), *Drosophila melanogaster* (fruit fly) and *Arabidopsis thaliana* (Arabidopsis, mouse-ear cress). In fact, one of the definitions of a model organism used to encompass (among other attributes) a small and simple genome. The main obstacle in whole-genome sequencing was the prohibitively high cost driven by the relatively low throughput of the traditional Sanger sequencing.

Sanger sequencing is based on dideoxynucleotide chemistries, with one read per reaction. The throughput and ease of Sanger was significantly increased through the introduction of capillary columns and the automation of the process

by Applied Biosystems (Martingallardo et al. 1992), but the basic shortcoming of the process – one read per reaction – remained unchanged and imposed a ceiling to the throughput of this technology. The understanding of this limitation led to the realization that a quantum leap in the process is necessary to meet the ever-increasing glut for sequence information. A massive search for new sequencing technologies resulted in several conceptually new sequencing methods. The majority of these methods employ the concept of massive, parallel sequencing of arrayed DNA fragments simultaneously using various light emitting chemistries (Mardis 2008, Metzker 2010) Thus, instead of sequencing one DNA template per reaction, tens of millions of DNA fragments could now be sequenced at the same time. These technologies are now known as next-generation sequencing (next-gen). The evolution of new sequencing methods is still ongoing. New technologies are coming online every year increasing sequencing accuracy, length and throughput, with seemingly no limits to the further improvement of the process.

The power of these technologies has brought democratization of genomics. Once an attribute and privilege of elite of a few species, today whole genome sequences are available for more than 180 eukaryotic organisms and genomic information is accumulating with unprecedented rate (Pagani et al. 2012). For example, the Joint Genomic Institute (JGI) of the US Department of Energy (DOE) has sequenced more than 40 plant genomes (phytozome.net). All of these are now easily and publicly available at the phytozome.net site, with powerful tools to compare genomics information among various linages of the plant evolutionary tree. In addition, genome sequencing is now expanding from the species to the individual level. A number of projects have been launched to sequence thousands of individual genomes from one species bringing genomics at the population level scales (Altshuler et al. 2012, Cao et al. 2011, Gan et al. 2011). This would enable unprecedented genome-based approaches in breeding, population genetics and conservation.

Towards Systems Biology Understanding through Various 'Omics' Approaches

The ultimate purpose of genomics is to facilitate functional understanding and thus its eventual usefulness in various fields and endeavors (Fig. 1). As we start to rapidly realize that most biological processes are highly interconnected networks involving large numbers of genes, it is becoming evident that understanding the function of the genome would entail understanding the function of genes *en block*, as a system, rather than as isolated entities. This has spurred interest in what is now known as other 'omics' approaches. Most prominently and following the central dogma, these include transcriptomics, proteomics and metabolomics focusing on RNA, proteins and metabolites respectively.

Transcriptomics involves approaches for analyses of genome-wide gene expression. Methods for studying gene expression have been around for a long time (e.g., Northern blots and RT-PCR). However they are suitable for studying expression of one or a few genes. Monitoring the expression of every single gene in the genome is a daunting task.

The first breakthrough was the development of microarrays (Schena et al. 1995). This technology uses a nucleic hybridization approach. DNA probes, representing portions of the genes in the genome, are arrayed on solid surface. The RNA is converted into cDNA using reverse transcriptase, labeled with fluorescent dyes and, allowed to hybridize to the array, followed by washing to remove unspecific (non-homologous) hybridizations. The amount of fluorescence at each probe is proportional to the labeled cDNA that hybridized to the probe and thus provides an indirect measure of the quantity of the respective mRNA (gene expression) in the sample of interest. The advances of next-generation sequencing are quickly revolutionizing these fields through methods now known as RNA sequencing (RNA-seq, Malone and Oliver 2011, Wang et al. 2009). As with microarrays, mRNA is reverse-transcribed into cDNA, which is then intensively sequenced using one of the above mentioned next-generation platforms. In contrast to the indirect methods of expression estimation using hybridization signal as a surrogate, these methods allow direct counting of transcript abundance and thus truly quantitative estimation of differences in gene expression. Furthermore, microarrays were limited to the number of features on the array and were also contingent on available genome sequence or large EST datasets. RNA-seq does not require either. The sensitivity and inferences of transcriptomics approaches have been greatly enhanced by using highly specific targeted sampling of tissues and even cell types using techniques like laser micro capture (LMC) (Larisch et al. 2012) and fluorescent activated cell sorting (FACS) (Birnbaum et al. 2003).

Proteomics emerged in the late nineties, closely following the rise of genomics and focusing on identification and quantification of protein abundance at a genome-wide scale (James 1997). Proteomics approaches have historically lagged behind transcriptomics due to the technological difficulties of studying proteins in a high throughput fashion. Initially, this was achieved through two-dimensional (2D) electrophoretic separation of protein mixtures, followed by spot isolation and sequencing through proteolytic degradation (Rabilloud and Lelong 2011, Wilm et al. 1996). However, these approaches lack the throughput and resolution that are necessary to scale up the method to levels approaching genome coverage (Vanderschuren et al. 2013). For example, even in the Arabidopsis model plant and most advanced analytical tools, only approximately an estimated 70 % coverage of the whole proteome has been achieved (Mann et al. 2013). Recently, mass spectroscopy using Orbitrap instruments and Multidimensional Protein Identification Technology (MudPIT) has emerged as the leading technology for ‘omics’ level protein analysis (Yates et al. 2009). The results from a controlled proteolysis are analyzed via extremely sensitive liquid chromatography/mass spectroscopy instruments that allow separation of thousands of proteins and accurate sequence assignment based on mass spectroscopic identification using mass and charge characteristics (Yates et al. 2009).

Similar to proteomics, metabolomics involves separation of complex metabolic mixtures by liquid or gas phase chromatographies (LC, GC) followed by mass spectroscopy (MS) or other chemical analytical procedures (Goodacre et al. 2004). These basic techniques have been com-

bined in various configurations to achieve better separation and identification of the analyzed mixtures (Putri et al. 2013). The bottleneck of metabolomics is the poor knowledge about many metabolites which, although well-detected as molecular entities, remain 'unknown' as chemical structures and are thus uninformative with respect to the underlying metabolic mechanisms.

In addition to understanding the various levels of information flows (e.g., DNA-RNA-protein-metabolite), some novel approaches attempt to understand the dynamic nature of DNA, particularly related to its regulation. These include the study of the epigenome where the DNA histone chemical modifications (e.g., methylation, acetylation) patterns are studied on a genome scale using bisulfide sequencing (Frommer et al. 1992) and chromatin immune precipitation (ChIP)-seq technologies (Barski et al. 2007, Mikkelsen et al. 2007). These techniques involve selective isolation of modified DNA variants and their subsequent sequencing using one of the next-gen platforms (Brinkman et al. 2012). This allows mapping of different modifications onto the genome space and associations of modification patterns with various factors (e.g., developmental and physiological phases). Such techniques provide a glimpse into a new type of polymorphisms, one that may act beyond the changes in nucleotide sequences but are rather imprinted through DNA modifications. This allows a more plastic and faster adjustment to rapidly changing conditions that does not require generations and cycles of natural selection to fix/enrich the most adaptive variants. The importance of epigenetic polymorphism to phenotypic diversity and its potential role in plant evolution have been recently shown (Becker et al. 2011, Paszkowski and Grossniklaus 2011, Schmitz et al. 2011). For example, promoter hypermethylation of the FLOWERING WAGENINGEN (FWA) gene in Arabidopsis delays transition to flowering and is maternally expressed likely through genetic imprinting (Fujimoto et al. 2008). Intriguingly, epigenetic states can be environmentally induced and thus could be of significant value to adaptation, particularly under fast-changing conditions (Lang-Mladek et al. 2010). An epigenetic mechanism contributing to fast climatic adaptation and involving microRNAs has been described in spruce (Yakovlev et al. 2010).

The ultimate goal of all these and other 'omics'-scale endeavors is to create the foundations for integrated, comprehensive system-based approaches to understanding biological processes (Long et al. 2008, Weckwerth 2011, Fig. 1). These approaches require new methods and predictive algorithms and models to accommodate the disparate nature of the data streams and the uncertainties about the underlying causative relations and mechanisms (Keurentjes et al. 2011, Saetzler et al. 2011).

Applications of Genomics in Forestry Research

Advances in genomics have been extremely empowering in many fields of research, including forestry. Early genomics efforts encompassed generation of EST resources in species of economic and environmental significance from the genera *Pinus*, *Picea*, *Eucalyptus*, *Quercus* and *Populus*. These efforts were tailored towards discovery of genes in various regulatory and metabolic processes of interest.

For example an intensive area of research was the wood-forming and particularly the lignin biosynthetic pathway. Although these first genomics projects did not extensively characterize the respective genomes, they provided important information about critical genes and enabled translation approaches based on homologies to model systems like Arabidopsis, where the processes were much better understood (Allona et al. 1998, Kirst et al. 2003). ESTs also empowered the generation of the first cDNA microarrays that allowed simultaneous interrogation of large numbers of genes involved in various processes in trees (Brinker et al. 2004, Hertzberg et al. 2001). The information generated through these EST projects in *Populus* was instrumental in the assembly and annotation of the first poplar tree genome (Tuskan et al. 2006). Finally, ESTs became a rich source for discovery of polymorphic markers that can be used in Quantitative Trait Loci (QTL) and physical mapping research (Liewlaksaneeyanawin et al. 2004, Varshney et al. 2005).

The collection of EST genomic information logically further escalated into whole-genome sequencing. This was first accomplished in *Populus* (Tuskan et al. 2006). The poplar genome was relatively small, well-mapped and abundant EST resources were available. The convergence of these factors allowed its successful sequencing, assembly and annotation. Since the originally published draft, the poplar genome has been updated several times to include some unresolved gaps and incorporate previously unmapped scaffolds (DNA fragments comprising tens of thousands of base pairs) into the genome sequence. The home of the poplar genome is the phytozome.net site, and it includes a genome browser and a rich resource for comparative genomics with other woody and herbaceous plants. The sequencing of the poplar genome is now followed by a number of other whole genome sequencing efforts of forest trees. These include eucalypt, pine and spruce (Mackay et al. 2012, Neale and Savolainen 2004, Nystedt et al. 2013). In addition, several species of trees and vines with importance to the fruit industry, like peach, apple and grapevine, were also sequenced (Jaillon et al. 2007, Velasco et al. 2010, Verde et al. 2013).

The genomics resources generated through these projects are empowering in several ways and have been tailored to facilitate breeding and biotechnological approaches to improve biomass yield, quality and resistance to biotic/abiotic stresses (Brinker et al. 2010, Guerra et al. 2013, Quesada et al. 2010, Ralph et al. 2006, Wegrzyn et al. 2010). Because sequencing projects were funded through public sources, the majority of this data is publicly available through online outlets (Table 1). The primary goal was and still is the detection of polymorphisms that can be used to assist breeding efforts by providing genotypic predictions to phenotypic variability in traits of economic and/or environmental significance. Initially, this was achieved using QTL and more recently association genetics (Brown et al. 2003, Gonzalez-Martinez et al. 2007, Quesada et al. 2010). The genome-wide availability of markers has also spurred interest in applying genome-wide selection approaches (Harfouche et al. 2012, Meuwissen et al. 2001, Resende M.D. et al. 2012, Resende M.F. et al. 2012)

In addition to the purely applied aspects of genomics research in breeding, there has been considerable interest in using genomics information for gaining more

fundamental understanding of processes that are typical for woody perennials and that are difficult or impossible to approach in herbaceous annual systems. These include wood formation, bud dormancy and some symbiotic interactions like mycorhizal associations (Brunner et al. 2004, Jansson and Douglas 2007). These studies are predominantly performed in poplar because of an efficient transformation system, which is not available or is difficult in any of the other trees species with significant genomics resources (Busov et al. 2005a, 2005b). Because of this, poplar has emerged as a 'model' taxon for understanding woody perennial development (Wullschleger et al. 2002). The importance of transformation as a functional tool in poplar has been reviewed elsewhere (Busov et al. 2005a, 2005b; Flachowsky et al. 2009). Briefly, the transformation can generate dominant lesions through using upregulation or gene-silencing mediated suppression of gene activity (e.g., RNAi or amiRNA). These dominant approaches bypass the long generation time and outcrossing system in trees, which makes performing traditional genetics, requiring rounds of selfing or close-relative mating to expose recessive mutations, an impractically long process. Several recent studies have demonstrated the power of these approaches in poplar (Bohlenius et al. 2006, Du et al. 2009, Hsu et al. 2006, Leple et al. 2007, Pilate et al. 2002, Plett et al. 2010, Rohde et al. 2002, Ruttink et al. 2007). In addition to validation function of known and/or candidate genes (reverse genetics), these approaches could also be used to find genes *de novo* through forward genetics by generating a population of T-DNA insertional mutants (Busov et al. 2010; Busov et al. 2005a, 2005b; Busov et al. 2003; Fladung et al. 2004; Fladung and Polak 2012; Harrison et al. 2007; Plett et al. 2010; Yordanov et al. 2010). The availability of genome sequence makes these approaches the fastest and most straightforward path to linking phenotypic to genotypic variation. The presence of known sequence (T-DNA) that affect gene function allows rapid recovery of sequences flanking the T-DNA and thus positioning the lesion in the genome sequence and providing the link between phenotypic changes and their genetic bases. The combination of transgenic manipulations with other 'omics' approaches, like for example transcriptomics, allows unprecedented insight into the regulatory mechanisms that underlie diverse biological processes and traits (Fig. 2).

The use of genomics has also recently been extended outside the traditional genetics area, into the ecology research arena. Genomics approaches in ecology have been known as metagenomics (beyond genomics) (Riesenfeld et al. 2004, Singh et al. 2009, Venter et al. 2004). This is an extremely promising and yet still under-employed area of research. In a nutshell, the presence of DNA from specific organism is indicative of its occurrence in a particular ecosystem (Simon and Daniel 2011). This is particularly interesting in elucidating the structure of microbial communities, which in many instances are impossible to study otherwise because cannot be readily cultured (Schloss and Handelsman 2005) and thus their presence and abundance is difficult to impossible to track with conventional culturing methods. These approaches not only provide estimates of presence/absence but also can quantify the composition, as well as shifts in the microbial communities in response to various experimental factors. Furthermore, RNA-seq (see above) or proteomics approaches applied to the same systems can

go a step further and not only identify the microbes in a sample but also what they are engaged in with respect to, for example, metabolic activities (Carvalhais et al. 2012, Simon and Daniel 2011, Wilmes and Bond 2006). This provides unprecedented insights into the function of the ecosystems at a new level that has been previously nearly ignored because of the lack of tools to approach it. Several recent studies have shown the power of this approach. A study in the root system of *Arabidopsis* plants has shown that specific soil samples drive particular endophytic communities associated with these samples (Lundberg et al. 2012). The development of these communities showed predictable and repeatable patterns, suggesting the existence of established and evolutionary conserved microbial-plant interactions.

The Road Ahead

One of the grand challenges in the 'post-genomic' area is the integration of all data streams into biological coherent information. This will require novel approaches for analysis, visualization and modeling. Ultimately, these integrative approaches will result with what is now known as virtual plant, best exemplified in the iPlant project (http://www.iplantcollaborative.org/). An even further-reaching project known as KBase (http://kbase.science.energy.gov/) is attempting to extend these integrative approaches to ecosystem scales. This will require novel software and analytical tools but also unprecedented computational power that can accommodate the demands of the analyses.

Enormous technological steps have made genomics, transcriptomics, proteomics and metabolomics approaches that are scalable to full genome scale. In stark contrast, and despite its prominent role in the system biology approaches (Fig. 2), genetics has remained a low throughput and generally a bottleneck. Therefore major technological breakthroughs are needed so that genetics measures to the power and throughput of the other 'omics'

Table 1. Publicly available genomics resources for various tree species.

Species/Genera/Family	Web site
Various european conifer and decidious tree species	http://www.evoltree.eu/index.php/e-resources
Quercus spp.	https://w3.pierroton.inra.fr/QuercusPortal/index.php?p=OAK_GENOME_SEQUENCING
Eucalyptus spp.	http://web.up.ac.za/eucagen/
Various hardwood tree species of North America	http://www.hardwoodgenomics.org/
Tree species of Fagaceae family	http://www.fagaceae.org/home
Populus spp.	http://popgenie.org/
Populus spp.	http://treesbio.com/
Tree and herabceous species of Rosaceae family	http://www.rosaceae.org/
Various conifer species and *Populus* spp.	http://www.arborea.ulaval.ca/
Various conifer species	http://www.treenomix.ca/
Pinus taeda	http://pinegenome.org/pinerefseq/
Various conifer species	http://congenie.org/

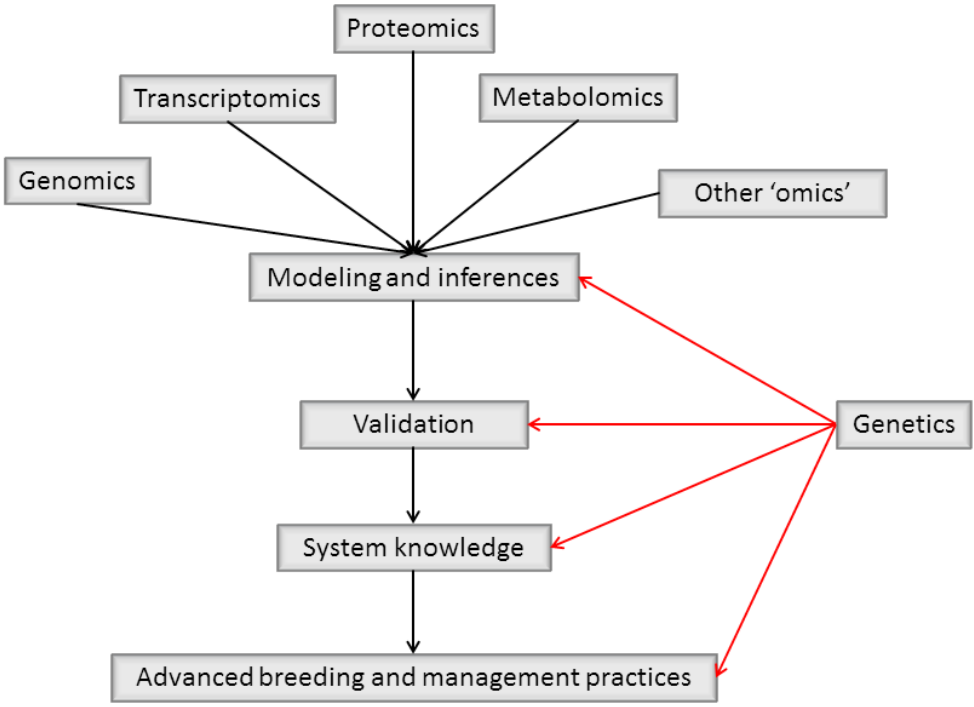

Fig. 1. Role of 'omics' approaches in generation of system level knowledge about biological systems.

Note the central place of genetics (indicated by red arrows) at various levels of generating, validating and implementing system knowledge. Genetics indicate both traditional and transgenic approaches for genetic analyses.

technologies. This is particularly important in the generation strong gain and loss-of-function mutations that are most effective in elucidation the gene function. Among plant species, it is only Arabidopsis that has mutagenized populations that approach genome coverage (Alonso et al. 2003). Such populations are not feasible in many species and thus require innovative approaches to generate viable alternatives. Some innovations involving zinc finger nucleases show promise but their utility across many species remains to be proven and are by no means are approaching even moderate throughout scale (Zhang et al. 2010).

Ultimately and most importantly, genomics has to yet decisively enter real mainstream life. Despite the technological democratization, genomics is still a privilege of science laboratories that can afford monetarily and intellectually the demands of such investigations. Although, the cost per base pair has been steadily and dramatically decreasing, these costs are based on high sample volumes requiring significant investments worth thou-

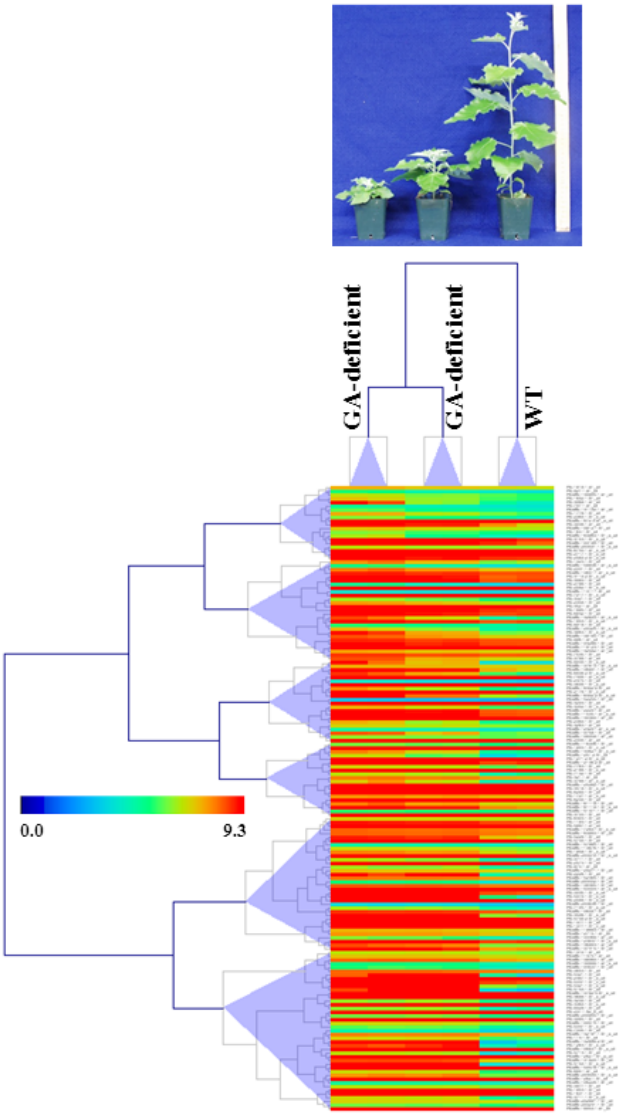

Fig. 2. Transgenics and 'omics' approaches provide a powerful system for molecular dissection of traits and processes.

GA-deficient and insensitive transgenic poplars are shown on the picture (Fig. 2). Note the characteristic dwarf phenotype of the GA-modified plants (left and middle) when compared to WT (right). A total of 1600 differentially regulated genes were found in the GA-insensitive and deficient poplars when compared to WT. Generation and characterization of the transgenics was previously described (Busov et al. 2006; Busov et al. 2003; Elias et al. 2012; Gou et al. 2010; Zawaski et al. 2011a, 2011b). The figure shows 168 upregulated

genes in the leaves of the GA-deficient and GA-insensitive poplar transgenics compared to WT. The picture at the top shows representative plants of each genotypic class as specified at the branches of the cluster and positioned above the respective columns showing the differentially regulated genes in each genotype. The figure was generated using hierarchical clustering. Colors indicate levels of expression as shown in the scale bar. Blue highlights indicate groups of genes and/or genotypes. WT=wild type.

sands of dollars. Therefore genomics has yet to become intellectually and economically approachable. This would require better genomics education, data sharing and significant expansion of 'genomics' service industry. In addition, genomics faces significant ethical problems related to intellectual property rights and the patentability of sequence data. Therefore one of the grand challenges in the future is making genomics and other 'omics' approaches intellectually, economically and sociably feasible and acceptable.

References

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies A., de Pablos B., Delcher A., Deng Z.M., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F.C., Gorrell J.H., Gu Z.P., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston D.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z.X., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.W., Lasko P., Lei Y.D., Levitsky A.A., Li J.Y., Li Z.Y., Liang Y., Lin X.Y., Liu X.J., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H.H., Wang X., Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G.G., Zhao Q., Zheng L.S., Zheng X.Q.H., Zhong F.N., Zhong W.Y., Zhou X.J., Zhu S.P., Zhu X.H., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C. 2000. The genome sequence of *Drosophila melanogaster.* Science 287: 2185–2195.

Adams M.D., Kelley J.M., Gocayne J.D., Dubnick M., Polymeropoulos M.H., Xiao H., Merril C.R., Wu A., Olde B., Moreno R.F., Kerlavage A.R., Mccombie W.R., Venter J.C.

1991. Complementary-DNA Sequencing – Expressed Sequence Tags and Human Genome Project. Science 252: 1651–1656.

Allona I., Quinn M., Shoop E., Swope K., St Cyr S., Carlis J., Riedl J., Retzel E., Campbell M.M., Sederoff R., Whetten R.W. 1998. Analysis of xylem formation in pine by cDNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 95: 9693–9698.

Alonso J.M., Stepanova A.N., Leisse T.J., Kim C.J., Chen H., Shinn P., Stevenson D.K., Zimmerman J., Barajas P., Cheuk R., Gadrinab C., Heller C., Jeske A., Koesema E., Meyers C.C., Parker H., Prednis L., Ansari Y., Choy N., Deen H., Geralt M., Hazari N., Hom E., Karnes M., Mulholland C., Ndubaku R., Schmidt I., Guzman P., guilar-Henonin L., Schmid M., Weigel D., Carter D.E., Marchand T., Risseeuw E., Brogden D., Zeko A., Crosby W.L., Berry C.C., Ecker J.R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301: 653–657.

Altshuler D.M., Durbin R.M., Abecasis G.R., Bentley D.R., Chakravarti A., Clark A.G., Donnelly P., Eichler E.E., Flicek P., Gabriel S.B., Gibbs R.A., Green E.D., Hurles M.E., Knoppers B.M., Korbel J.O., Lander E.S., Lee C., Lehrach H., Mardis E.R., Marth G.T., McVean G.A., Nickerson D.A., Schmidt J.P., Sherry S.T., Wang J., Wilson R.K., Gibbs R.A., Dinh H., Kovar C., Lee S., Lewis L., Muzny D., Reid J., Wang M., Wang J., Fang X.D., Guo X.S., Jian M., Jiang H., Jin X., Li G.Q., Li J.X., Li Y.R., Li Z., Liu X., Lu Y., Ma X.D., Su Z., Tai S.S., Tang M.F., Wang B., Wang G.B., Wu H.L., Wu R.H., Yin Y., Zhang W.W., Zhao J., Zhao M.R., Zheng X.L., Zhou Y., Gabriel S.B., Gupta N., Flicek P., Clarke L., Leinonen R., Smith R.E., Zheng-Bradley X., Grocock R., Humphray S., James T., Kingsbury Z., Lehrach H., Sudbrak R., Albrecht M.W., Amstislavskiy V.S., Borodina T.A., Lienhard M., Mertes F., Sultan M., Timmermann B., Yaspo M.L., Sherry S.T., McVean G.A., Mardis E.R., Wilson R.K., Fulton L., Fulton R., Weinstock G.M., Balasubramaniam S., Burton J., Danecek P., Keane T.M., Kolb-Kokocinski A., McCarthy S., Stalker J., Quail M., Schmidt J.P., Davies C.J., Gollub J., Webster T., Wong B., Zhan Y.P., Auton A., Gibbs R.A., Yu F., Bainbridge M., Challis D., Evani U.S., Lu J., Muzny D., Nagaswamy U., Reid J., Sabo A., Wang Y., Yu J., Wang J., Coin L.J.M., Fang L., Guo X.S., Jin X., Li G.Q., Li Q.B., Li Y.R., Li Z.Y., Lin H.X., Liu B.H., Luo R.B., Qin N., Shao H.J., Wang B.Q., Xie Y.L., Ye C., Yu C., Zhang F., Zheng H.C., Zhu H.M., Marth G.T., Garrison E.P., Kural D., Lee W.P., Leong W.F., Ward A.N., Wu J.T., Zhang M.Y., Lee C., Griffin L., Hsieh C.H., Mills R.E., Shi X.H., von Grotthuss M., Zhang C.S., Daly M.J., DePristo M.A., Altshuler D.M., Banks E., Bhatia G., Carneiro M.O., del Angel G., Gabriel S.B., Genovese G., Gupta N., Handsaker R.E., Hartl C., McCarroll S.A., Nemesh J.C., Poplin R.E., Schaffner S.F., Shakir K., Yoon S.C., Lihm J., Makarov V., Jin H.J., Kim W., Kim K.C., Korbel J.O., Rausch T., Flicek P., Beal K., Cunningham F., Herrero J., McLaren W.M., Ritchie G.R.S., Smith R.E., Zheng-Bradley X., Gottipati S., Keinan A., Rodriguez-Flores J.L., Sabeti P.C., Grossman S.R., Tabrizi S., Tariyal R., Cooper D.N., Ball E.V., Stenson P.D., Barnes B., Bauer M., Cheetham R.K., Cox T., Eberle M., Humphray S., Kahn S., Murray L., Peden J., Shaw R., Ye K., Batzer M.A., Konkel M.K., Walker J.A., MacArthur D.G., Lek M., Sudbrak R., Amstislavskiy V.S., Herwig R., Shriver M.D., Bustamante C.D., Byrnes J.K., De la Vega F.M., Gravel S., Kenny E.E., Kidd J.M., Lacroute P., Maples B.K., Moreno-Estrada A., Zakharia F., Halperin E., Baran Y., Craig D.W., Christoforides A., Homer N., Izatt T., Kurdoglu A.A., Sinari S.A., Squire K., Sherry S.T., Xiao C.L., Sebat J., Bafna V., Ye K., Burchard E.G., Hernandez R.D., Gignoux C.R., Haussler D., Katzman S.J., Kent W.J., Howie B., Ruiz-Linares A., Dermitzakis E.T., Lappalainen T., Devine S.E.,

Liu X.Y., Maroo A., Tallon L.J., Rosenfeld J.A., Michelson L.P., Kang H.M., Anderson P., Angius A., Bigham A., Blackwell T., Busonero F., Cucca F., Fuchsberger C., Jones C., Jun G., Li Y., Lyons R., Maschio A., Porcu E., Reinier F., Sanna S., Schlessinger D., Sidore C., Tan A., Trost M.K., Awadalla P., Hodgkinson A., Lunter G., McVean G.A., Marchini J.L., Myers S., Churchhouse C., Delaneau O., Gupta-Hinch A., Iqbal Z., Mathieson I., Rimmer A., Xifara D.K., Oleksyk T.K., Fu Y.X., Liu X.M., Xiong M.M., Jorde L., Witherspoon D., Xing J.C., Eichler E.E., Browning B.L., Alkan C., Hajirasouliha I., Hormozdiari F., Ko A., Sudmant P.H., Mardis E.R., Chen K., Chinwalla A., Ding L., Dooling D., Koboldt D.C., McLellan M.D., Wallis J.W., Wendl M.C., Zhang Q.Y., Durbin R.M., Hurles M.E., Tyler-Smith C., Albers C.A., Ayub Q., Balasubramaniam S., Chen Y., Coffey A.J., Colonna V., Danecek P., Huang N., Jostins L., Keane T.M., Li H., McCarthy S., Scally A., Stalker J., Walter K., Xue Y.L., Zhang Y.J., Gerstein M.B., Abyzov A., Balasubramanian S., Chen J.M., Clarke D., Fu Y., Habegger L., Harmanci A.O., Jin M.K., Khurana E., Mu X.J., Sisu C., Li Y.R., Luo R.B., Zhu H.M., Lee C., Griffin L., Hsieh C.H., Mills R.E., Shi X.H., von Grotthuss M., Zhang C.S., Marth G.T., Garrison E.P., Kural D., Lee W.P., Ward A.N., Wu J.T., Zhang M.Y., McCarroll S.A., Altshuler D.M., Banks E., del Angel G., Genovese G., Handsaker R.E., Hartl C., Nemesh J.C., Shakir K., Yoon S.C., Lihm J., Makarov V., Degenhardt J., Flicek P., Smith R.E., Zheng-Bradley X., Korbel J.O., Rausch T., Stutz A.M., Barnes B., Cheetham R.K., Eberle M., Humphray S., Kahn S., Murray L., Shaw R., Ye K., Batzer M.A., Konkel M.K., Walker J.A., Lacroute P., Craig D.W., Homer N., Church D., Xiao C.L., Sebat J., Bafna V., Michaelson J.J., Ye K., Devine S.E., Liu X.Y., Maroo A., Tallon L.J., Lunter G., McVean G.A., Iqbal Z., Witherspoon D., Xing J.C., Eichler E.E., Alkan C., Hajirasouliha I., Hormozdiari F., Ko A., Sudmant P.H., Chen K., Chinwalla A., Ding L., McLellan M.D., Wallis J.W., Hurles M.E., Blackburne B., Li H., Lindsay S.J., Ning Z.M., Scally A., Walter K., Zhang Y.J., Gerstein M.B., Abyzov A., Chen J.M., Clarke D., Khurana E., Mu X.J., Sisu C., Gibbs R.A., Yu F.L., Bainbridge M., Challis D., Evani U.S., Kovar C., Lewis L., Lu J., Muzny D., Nagaswamy U., Reid J., Sabo A., Yu J., Guo X.S., Li Y.R., Wu R.H., Marth G.T., Garrison E.P., Leong W.F., Ward A.N., del Angel G., DePristo M.A., Gabriel S.B., Gupta N., Hartl C., Poplin R.E., Clark A.G., Rodriguez-Flores J.L., Flicek P., Smith R.E., Zheng-Bradley X., MacArthur D.G., Bustamante C.D., Gravel S., Craig D.W., Christoforides A., Homer N., Izatt T., Sherry S.T., Xiao C.L., Dermitzakis E.T., Kang H.M., McVean G.A., Mardis E.R., Dooling D., Fulton L., Fulton R., Koboldt D.C., Durbin R.M., Balasubramaniam S., Keane T.M., McCarthy S., Stalker J., Gerstein M.B., Balasubramanian S., Habegger L., Garrison E.P., Gibbs R.A., Bainbridge M., Muzny D., Yu F.L., Yu J., del Angel G., Handsaker R.E., Makarov V., Rodriguez-Flores J.L., Jin H.J., Kim W., Kim K.C., Flicek P., Beal K., Cunningham F., Herrero J., McLaren W.M., Ritchie G.R.S., Zheng-Bradley X., Tabrizi S., MacArthur D.G., Lek M., Bustamante C.D., De la Vega F.M., Craig D.W., Kurdoglu A.A., Lappalainen T., Rosenfeld J.A., Michelson L.P., Awadalla P., Hodgkinson A., McVean G.A., Chen K., Tyler-Smith C., Chen Y., Colonna V., Frankish A., Harrow J., Xue Y.L., Gerstein M.B., Abyzov A., Balasubramanian S., Chen J.M., Clarke D., Fu Y., Harmanci A.O., Jin M.K., Khurana E., Mu X.J., Sisu C., Gibbs R.A., Fowler G., Hale W., Kalra D., Kovar C., Muzny D., Reid J., Wang J., Guo X., Li G., Li Y., Zheng X., Altshuler D.M., Flicek P., Barker J., Kelman G., Kulesha E., Leinonen R., McLaren W.M., Radhakrishnan R., Roa A., Smirnov D., Smith R.E., Streeter I., Toneva I., Vaughan B., Zheng-Bradley X., Cox T., Humphray S., Kahn S., Sudbrak R., Albrecht M.W., Lienhard M., Craig D.W., Izatt T.,

Kurdoglu A.A., Sherry S.T., Ananiev V., Belaia Z., Beloslyudtsev D., Bouk N., Chen C., Church D., Cohen R., Cook C., Garner J., Hefferon T., Kimelman M., Liu C., Lopez J., Meric P., O'Sullivan C., Ostapchuk Y., Phan L., Ponomarov S., Schneider V., Shekhtman E., Sirotkin K., Slotta D., Xiao C.L., Zhang H., Haussler D., McVean G.A., Alkan C., Ko A., Dooling D., Durbin R.M., Balasubramaniam S., Keane T.M., McCarthy S., Stalker J., Chakravarti A., Knoppers B.M., Barnes K.C., Beiswanger C., Burchard E.G., Bustamante C.D., Cai H.Y., Cao H.Z., Durbin R.M., Gharani N., Gibbs R.A., Gignoux C.R., Gravel S., Henn B., Jones D., Jorde L., Kaye J.S., Keinan A., Kent A., Kerasidou A., Li Y.R., Mathias R., McVean G.A., Moreno-Estrada A., Ossorio P.N., Parker M., Reich D., Rotimi C.N., Royal C.D., Sandoval K., Su Y.Y., Sudbrak R., Tian Z.M., Timmermann B., Tishkoff S., Toji L.H., Tyler-Smith C., Via M., Wang Y.H., Yang H.M., Yang L., Zhu J.Y., Bodmer W., Bedoya G., Ruiz-Linares A., Ming C.Z., Yang G., You C.J., Peltonen L., Garcia-Montero A., Orfao A., Dutil J., Martinez-Cruzado J.C., Oleksyk T.K., Brooks L.D., Felsenfeld A.L., McEwen J.E., Clemm N.C., Duncanson A., Dunn M., Green E.D., Guyer M.S., Peterson J.L., Auton A., Brooks L.D., DePristo M.A., Durbin R.M., Handsaker R.E., Kang H.M., Marth G.T., McVean G.A., Consortium G.P. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56–65.

Barski A., Cuddapah S., Cui K., Roh T.Y., Schones D.E., Wang Z., Wei G., Chepelev I., Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129: 823–837.

Becker C., Hagmann J., Muller J., Koenig D., Stegle O., Borgwardt K., Weigel D. 2011. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature 480: 245–249

Fujimoto R., Kinoshita Y., Kawabe A., Kinoshita T., Takashima K., Nordborg M., Nasrallah M.E., Shimizu K.K., Kudoh H., Kakutani T. 2008. Evolution and control of imprinted FWA genes in the genus *Arabidopsis*. PLoS genetics 4:e1000048

Birnbaum K., Shasha D.E., Wang J.Y., Jung J.W., Lambert G.M., Galbraith D.W., Benfey P.N. 2003. A gene expression map of the *Arabidopsis* root. Science 302: 1956–1960.

Blattner F.R., Plunkett G., Bloch C.A., Perna N.T., Burland V., Riley M., ColladoVides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277: 1453–&.

Bohlenius H., Huang T., Charbonnel-Campaa L., Brunner A.M., Jansson S., Strauss S.H., Nilsson O. 2006. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science 312: 1040–1043.

Brinker M., Brosche M., Vinocur B., Abo-Ogiala A., Fayyaz P., Janz D., Ottow E.A., Cullmann A.D., Saborowski J., Kangasjarvi J., Altman A., Polle A. 2010. Linking the Salt Transcriptome with Physiological Responses of a Salt-Resistant *Populus* Species as a Strategy to Identify Genes Important for Stress Acclimation. Plant Physiologyogy 154: 1697–1709.

Brinker M., van Z.L., Liu W., Craig D., Sederoff R.R., Clapham D.H., von A.S. 2004. Microarray analyses of gene expression during adventitious root development in *Pinus contorta*. Plant Physiology 135: 1526–1539.

Brinkman A.B., Gu H.C., Bartels S.J.J., Zhang Y.Y., Matarese F., Simmer F., Marks H., Bock C., Gnirke A., Meissner A., Stunnenberg H.G. 2012. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Research 22: 1128–1138.

Brown G.R., Bassoni D.L., Gill G.P., Fontana J.R., Wheeler N.C., Megraw R.A., Davis M.F., Sewell M.M., Tuskan G.A., Neale D.B. 2003. Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL Verification and candidate gene mapping. Genetics 164: 1537–1546.

Brunner A.M., Busov V.B., Strauss S.H. 2004. The Poplar genome sequence: Functional genomics in a keystone plant species. Trends in Plant Science 9: 49–56.

Busov V., Meilan R., Pearce D.W., Rood S.B., Ma C.P., Tschaplinski T.J., Strauss S.H. 2006. Transgenic modification of gai or rgl1 causes dwarfing and alters gibberellins, root growth, and metabolite profiles in *Populus*. Planta 224: 288–299.

Busov V., Yordanov Y., Gou J., Meilan R., Ma C., Regan S., Strauss S. 2010. Activation tagging is an effective gene tagging system in *Populus*. Tree Genetics and Genomes 7: 91–101.

Busov V.B., Brunner A.M., Meilan R., Filichkin S., Ganio L., Gandhi S., Strauss S.H. 2005a. Genetic transformation: a powerful tool for dissection of adaptive traits in trees. New Phytologist 167: 9–18.

Busov V.B., Fladung M., Groover A., Strauss S.H. 2005b. Insertional mutagenesis in *Populus*: relevance and feasibility. Tree Genetics and Genomes 1: 135–142.

Busov V.B., Meilan R., Pearce D.W., Ma C., Rood S.B., Strauss S.H. 2003. Activation tagging of a dominant gibberellin catabolism gene (GA 2-oxidase) from poplar that regulates tree stature. Plant Physiology 132: 1283–1291.

Cao J., Schneeberger K., Ossowski S., Gunther T., Bender S., Fitz J., Koenig D., Lanz C., Stegle O., Lippert C., Wang X., Ott F., Muller J., Alonso-Blanco C., Borgwardt K., Schmid K.J., Weigel D. 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nature Genetics 43: 956–960.

Carvalhais L.C., Dennis P.G., Tyson G.W., Schenk P.M. 2012. Application of metatranscriptomics to soil environments. Journal of Microbiology Methods 91: 246–251.

C. elegans sequencing consortium. 1998. Genome sequence of the nematode *C. elegans:* A platform for investigating biology. Science 282: 2012–2018.

Du J., Mansfield S.D., Groover A.T. 2009. The *Populus* homeobox gene ARBORKNOX2 regulates cell differentiation during secondary growth. Plant Journal 60: 1000–1014.

Elias A.A., Busov V.B., Kosola K.R., Ma C., Etherington E., Shevchenko O., Gandhi H., Pearce D.W., Rood S.B., Strauss S.H. 2012. Green revolution trees: semidwarfism transgenes modify gibberellins, promote root growth, enhance morphological diversity, and reduce competitiveness in hybrid poplar. Plant Physiology 160: 1130–1144.

Flachowsky H., Hanke M.V., Peil A., Strauss S.H., Fladung M. 2009. A review on transgenic approaches to accelerate breeding of woody plants. Plant Breeding 128: 217–226.

Fladung M., Deutsch F., Honicka H., Kumar S. 2004. T-DNA and transposon tagging in aspen. Plant Biol (Stuttg) 6: 5–11.

Fladung M., Polak O. 2012. Ac/Ds-transposon activation tagging in poplar: a powerful tool for gene discovery. BMC Genomics 13: 61.

Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M., Mckenney K., Sutton G., Fitzhugh W., Fields C., Gocayne J.D., Scott J., Shirley R., Liu L.I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M., Gnehm C.L., Mcdonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C. 1995. Whole-Genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269: 496–512.

Frommer M., Mcdonald L.E., Millar D.S., Collis C.M., Watt F., Grigg G.W., Molloy P.L., Paul C.L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proceedings of the National Academy of Sciences of the United States of America 89: 1827–1831.

Fujimoto R., Kinoshita Y., Kawabe A., Kinoshita T., Takashima K., Nordborg M., Nasrallah M.E., Shimizu K.K., Kudoh H., Kakutani T. 2008. Evolution and control of imprinted FWA genes in the genus *Arabidopsis*. PLoS Genetics 4:e1000048

Gan X.C., Stegle O., Behr J., Steffen J.G., Drewe P., Hildebrand K.L., Lyngsoe R.,

Schultheiss S.J., Osborne E.J., Sreedharan V.T., Kahles A., Bohnert R., Jean G., Derwent P., Kersey P., Belfield E.J., Harberd N.P., Kemen E., Toomajian C., Kover P.X., Clark R.M., Ratsch G., Mott R. 2011. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature 477: 419–423.

Goff S.A., Ricke D., Lan T.H., Presting G., Wang R.L., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H., Hadley D., Hutchinson D., Martin C., Katagiri F., Lange B.M., Moughamer T., Xia Y., Budworth P., Zhong J.P., Miguel T., Paszkowski U., Zhang S.P., Colbert M., Sun W.L., Chen L.L., Cooper B., Park S., Wood T.C., Mao L., Quail P., Wing R., Dean R., Yu Y.S., Zharkikh A., Shen R., Sahasrabudhe S., Thomas A., Cannings R., Gutin A., Pruss D., Reid J., Tavtigian S., Mitchell J., Eldredge G., Scholl T., Miller R.M., Bhatnagar S., Adey N., Rubano T., Tusneem N., Robinson R., Feldhaus J., Macalma T., Oliphant A., Briggs S. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296: 92–100.

Goffeau A., Barrell B.G., Bussey H., Davis R.W., Dujon B., Feldmann H., Galibert F., Hoheisel J.D., Jacq C., Johnston M., Louis E.J., Mewes H.W., Murakami Y., Philippsen P., Tettelin H., Oliver S.G. 1996. Life with 6000 genes. Science 274: 546, 563–547.

Gonzalez-Martinez S.C., Wheeler N.C., Ersoz E., Nelson C.D., Neale D.B. 2007. Association genetics in *Pinus taeda* L. I. Wood property traits. Genetics 175: 399–409.

Goodacre R., Vaidyanathan S., Dunn W.B., Harrigan G.G., Kell D.B. 2004. Metabolomics by numbers: acquiring and understanding global metabolite data. Trends in Biotechnology 22: 245–252.

Gou J., Strauss S.H., Tsai C.J., Fang K., Chen Y., Jiang X., Busov V.B. 2010. Gibberellins regulate lateral root formation in *Populus* through interactions with auxin and other hormones. Plant Cell 22: 623–639.

Guerra F.P., Wegrzyn J.L., Sykes R., Davis M.F., Stanton B.J., Neale D.B. 2013. Association genetics of chemical wood properties in black poplar (*Populus nigra*). New Phytologist 197: 162–176.

Harfouche A., Meilan R., Kirst M., Morgante M., Boerjan W., Sabatti M., Scarascia Mugnozza G. 2012. Accelerating the domestication of forest trees in a changing world. Trends in Plant Science 17: 64–72.

Harrison E.J., Bush M., Plett J.M., McPhee D.P., Vitez R., O'Malley B., Sharma V., Bosnich W., Seguin A., MacKay J., Regan S. 2007. Diverse developmental mutants revealed in an activation-tagged population of poplar. Canadian Journal of Botany 85: 1071–1081.

Hertzberg M., Aspeborg H., Schrader J., Andersson A., Erlandsson R., Blomqvist K., Bhalerao R., Uhlen M., Teeri T.T., Lundeberg J., Sundberg B., Nilsson P., Sandberg G. 2001. A transcriptional roadmap to wood formation. Proceedings of the National Academy of Sciences of the United States of America 98: 14732–14737.

Hsu C.Y., Liu Y., Luthe D.S., Yuceer C. 2006. Poplar FT2 shortens the juvenile phase and promotes seasonal flowering. Plant Cell 18: 1846–1861.

Jackson D.A., Symons R.H., Berg P. 1972. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 69: 2904–2909.

Jaillon O., Aury J.M., Noel B., Policriti A., Clepet C., Casagrande A., Choisne N., Aubourg S., Vitulo N., Jubin C., Vezzi A., Legeai F., Hugueney P., Dasilva C., Horner D., Mica E., Jublot D., Poulain J., Bruyere C., Billault A., Segurens B., Gouyvenoux M., Ugarte E., Cattonaro F., Anthouard V., Vico V., Del F.C., Alaux M., Di G.G., Dumas V., Felice N., Paillard S., Juman I., Moroldo M., Scalabrin S., Canaguier A., Le C.I., Malacrida G., Durand E., Pesole G., Laucou V., Chatelet P., Merdinoglu D., Delledonne M., Pezzotti M., Lecharny A., Scarpelli C., Artiguenave F., Pe M.E., Valle G., Morgante M., Caboche M., dam-Blondon A.F., Weissenbach J., Quetier F., Wincker P. 2007. The grapevine genome sequence suggests ancestral

hexaploidization in major angiosperm phyla. Nature 449: 463–467.

James P. 1997. Protein identification in the post-genome era: the rapid rise of proteomics. Quarterly Reviews of Biophysics 30: 279–331.

Jansson S., Douglas C.J. 2007. *Populus*: a model system for plant biology. Annual Review of Plant Biology 58: 435–458.

Kaul S., Koo H.L., Jenkins J., Rizzo M., Rooney T., Tallon L.J., Feldblyum T., Nierman W., Benito M.I., Lin X.Y., Town C.D., Venter J.C., Fraser C.M., Tabata S., Nakamura Y., Kaneko T., Sato S., Asamizu E., Kato T., Kotani H., Sasamoto S., Ecker J.R., Theologis A., Federspiel N.A., Palm C.J., Osborne B.I., Shinn P., Conway A.B., Vysotskaia V.S., Dewar K., Conn L., Lenz C.A., Kim C.J., Hansen N.F., Liu S.X., Buehler E., Altafi H., Sakano H., Dunn P., Lam B., Pham P.K., Chao Q., Nguyen M., Yu G.X., Chen H.M., Southwick A., Lee J.M., Miranda M., Toriumi M.J., Davis R.W., Wambutt R., Murphy G., Dusterhoft A., Stiekema W., Pohl T., Entian K.D., Terryn N., Volckaert G., Salanoubat M., Choisne N., Rieger M., Ansorge W., Unseld M., Fartmann B., Valle G., Artiguenave F., Weissenbach J., Quetier F., Wilson R.K., de la Bastide M., Sekhon M., Huang E., Spiegel L., Gnoj L., Pepin K., Murray J., Johnson D., Habermann K., Dedhia N., Parnell L., Preston R., Hillier L., Chen E., Marra M., Martienssen R., McCombie W.R., Mayer K., White O., Bevan M., Lemcke K., Creasy T.H., Bielke C., Haas B., Haase D., Maiti R., Rudd S., Peterson J., Schoof H., Frishman D., Morgenstern B., Zaccaria P., Ermolaeva M., Pertea M., Quackenbush J., Volfovsky N., Wu D.Y., Lowe T.M., Salzberg S.L., Mewes H.W., Rounsley S., Bush D., Subramaniam S., Levin I., Norris S., Schmidt R., Acarkan A., Bancroft I., Quetier F., Brennicke A., Eisen J.A., Bureau T., Legault B.A., Le Q.H., Agrawal N., Yu Z., Martienssen R., Copenhaver G.P., Luo S., Pikaard C.S., Preuss D., Paulsen I.T., Sussman M., Britt A.B., Selinger D.A., Pandey R., Mount D.W., Chandler V.L., Jorgensen R.A., Pikaard C., Juergens G., Meyerowitz E.M., Theologis A., Dangl J., Jones J.D.G., Chen M., Chory J., Somerville M.C., In A.G. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796–815.

Keurentjes J.J.B., Angenent G.C., Dicke M., Dos Santos V.A.P.M., Molenaar J., van der Putten W.H., de Ruiter P.C., Struik P.C., Thomma B.P.H.J. 2011. Redefining plant systems biology: from cell to ecosystem. Trends in Plant Science 16: 183–190.

Kirst M., Johnson A.F., 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 of the United States of America 100: 7383–7388.

Lander E.S., Consortium I.H.G.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J.P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J.C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R.H., Wilson R.K., Hillier L.W., McPherson J.D., Marra M.A., Mardis E.R., Fulton L.A., Chinwalla A.T., Pepin K.H., Gish W.R., Chissoe S.L., Wendl M.C., Delehaunty K.D., Miner T.L., Delehaunty A., Kramer J.B., Cook L.L., Fulton R.S., Johnson D.L., Minx P.J., Clifton S.W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J.F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., Gibbs R.A., Muzny D.M., Scherer S.E., Bouck J.B., Sodergren E.J., Worley K.C., Rives C.M., Gorrell J.H., Metzker M.L., Naylor S.L., Kucherlapati R.S., Nelson D.L., Weinstock G.M., Sakaki Y., Fujiyama A., Hattori M., Yada

T., Toyoda A., Itoh T., Kawagoe C., Watanabe H., Totoki Y., Taylor T., Weissenbach J., Heilig R., Saurin W., Artiguenave F., Brottier P., Bruls T., Pelletier E., Robert C., Wincker P., Rosenthal A., Platzer M., Nyakatura G., Taudien S., Rump A., Yang H.M., Yu J., Wang J., Huang G.Y., Gu J., Hood L., Rowen L., Madan A., Qin S.Z., Davis R.W., Federspiel N.A., Abola A.P., Proctor M.J., Myers R.M., Schmutz J., Dickson M., Grimwood J., Cox D.R., Olson M.V., Kaul R., Raymond C., Shimizu N., Kawasaki K., Minoshima S., Evans G.A., Athanasiou M., Schultz R., Roe B.A., Chen F., Pan H.Q., Ramser J., Lehrach H., Reinhardt R., McCombie W.R., de la Bastide M., Dedhia N., Blocker H., Hornischer K., Nordsiek G., Agarwala R., Aravind L., Bailey J.A., Bateman A., Batzoglou S., Birney E., Bork P., Brown D.G., Burge C.B., Cerutti L., Chen H.C., Church D., Clamp M., Copley R.R., Doerks T., Eddy S.R., Eichler E.E., Furey T.S., Galagan J., Gilbert J.G.R., Harmon C., Hayashizaki Y., Haussler D., Hermjakob H., Hokamp K., Jang W.H., Johnson L.S., Jones T.A., Kasif S., Kaspryzk A., Kennedy S., Kent W.J., Kitts P., Koonin E.V., Korf I., Kulp D., Lancet D., Lowe T.M., McLysaght A., Mikkelsen T., Moran J.V., Mulder N., Pollara V.J., Ponting C.P., Schuler G., Schultz J.R., Slater G., Smit A.F.A., Stupka E., Szustakowki J., Thierry-Mieg D., Thierry-Mieg J., Wagner L., Wallis J., Wheeler R., Williams A., Wolf Y.I., Wolfe K.H., Yang S.P., Yeh R.F., Collins F., Guyer M.S., Peterson J., Felsenfeld A., Wetterstrand K.A., Patrinos A., Morgan M.J., Conso I.H.G.S. 2001. Initial sequencing and analysis of the human genome. Nature 409: 860–921.

Lang-Mladek C., Popova O., Kiok K., Berlinger M., Rakic B., Aufsatz W., Jonak C., Hauser M.T., Luschnig C. 2010. Transgenerational inheritance and resetting of stress-induced loss of epigenetic gene silencing in *Arabidopsis*. Molecular plant 3: 594–602.

Larisch C., Dittrich M., Wildhagen H., Lautner S., Fromm J., Polle A., Hedrich R., Rennenberg H., Muller T., Ache P. 2012. Poplar wood rays are involved in seasonal remodeling of tree physiology. Plant Physiology 160: 1515–1529.

Leple J.C., Dauwe R., Morreel K., Storme V., Lapierre C., Pollet B., Naumann A., Kang K.Y., Kim H., Ruel K., Lefebvre A., Joseleau J.P., Grima-Pettenati J., De Rycke R., ndersson-Gunneras S., Erban A., Fehrle I., Petit-Conil M., Kopka J., Polle A., Messens E., Sundberg B., Mansfield S.D., Ralph J., Pilate G., Boerjan W. 2007. Downregulation of cinnamoyl-coenzyme a reductase in poplar: Multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. Plant Cell 19: 3669–3691.

Liewlaksaneeyanawin C., Ritland C.E., El-Kassaby Y.A., Ritland K. 2004. Single-copy, species-transferable microsatellite markers developed from Loblolly pine ESTs. Theoretical and Applied Genetics 109: 361–369.

Long T.A., Brady S.M., Benfey P.N. 2008. Systems approaches to identifying gene regulatory networks in plants. Annual Review of Cell Developmental Biology 24: 81–103.

Lundberg D.S., Lebeis S.L., Paredes S.H., Yourstone S., Gehring J., Malfatti S., Tremblay J., Engelbrektson A., Kunin V., del Rio T.G., Edgar R.C., Eickhorst T., Ley R.E., Hugenholtz P., Tringe S.G., Dangl J.L. 2012. Defining the core *Arabidopsis thaliana* root microbiome. Nature 488: 86–90.

Luria S.E., Human M.L. 1952. A nonhereditary, host-induced variation of bacterial viruses. Journal of bacteriology 64: 557–569.

Mackay J., Dean J.F., Plomion C., Peterson D.G., Canovas F.M., Pavy N., Ingvarsson P.K., Savolainen O., Guevara M.A., Fluch S., Vinceti B., Abarca D., Diaz-Sala C., Cervera M.T. 2012. Towards decoding the conifer giga-genome. Plant Molecular Biology 80: 555–569.

Malone J.H., Oliver B. 2011. Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biology 9(34): 1–9.

Mann G.W., Joshi H.J., Petzold C.J., Heazlewood J.L. 2013. Proteome coverage of the model plant *Arabidopsis thaliana*: implications for shotgun proteomic studies. Journal of Proteomics 79: 195–199.

Mardis E.R. 2008. Next-generation DNA sequencing methods. Annual Review of Genomics and Human Genetics 9: 387–402.

Martingallardo A., Mccombie W.R., Gocayne J.D., Fitzgerald M.G., Wallace S., Lee B.M.B., Lamerdin J., Trapp S., Kelley J.M., Liu L.I., Dubnick M., Johnstondow L.A., Kerlavage A.R., Dejong P., Carrano A., Fields C., Venter J.C. 1992. Automated DNA sequencing and analysis of 106 Kilobases from human chromosome-19q13.3. Nature Genetics1: 34–39.

Meuwissen T.H., Hayes B.J., Goddard M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.

Mikkelsen T.S., Ku M., Jaffe D.B., Issac B., Lieberman E., Giannoukos G., Alvarez P., Brockman W., Kim T.K., Koche R.P., Lee W., Mendenhall E., O'Donovan A., Presser A., Russ C., Xie X., Meissner A., Wernig M., Jaenisch R., Nusbaum C., Lander E.S., Bernstein B.E. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448: 553–560.

Neale D.B., Savolainen O. 2004. Association genetics of complex traits in conifers. Trends in Plant Science 9: 325–330.

Nystedt B., Street N.R., Wetterbom A., Zuccolo A., Lin Y.C., Scofield D.G., Vezzi F., Delhomme N., Giacomello S., Alexeyenko A., Vicedomini R., Sahlin K., Sherwood E., Elfstrand M., Gramzow L., Holmberg K., Hallman J., Keech O., Klasson L., Koriabine M., Kucukoglu M., Kaller M., Luthman J., Lysholm F., Niittyla T., Olson A., Rilakovic N., Ritland C., Rossello J.A., Sena J., Svensson T., Talavera-Lopez C., Theissen G., Tuominen H., Vanneste K., Wu Z.Q., Zhang B., Zerbe P., Arvestad L., Bhalerao R., Bohlmann J., Bousquet J., Garcia Gil R., Hvidsten T.R., de Jong P., MacKay J., Morgante M., Ritland K., Sundberg B., Thompson S.L., Van de Peer Y., Andersson B., Nilsson O., Ingvarsson P.K., Lundeberg J., Jansson S. 2013. The Norway spruce genome sequence and conifer genome evolution. Nature 497: 579–584.

Pagani I., Liolios K., Jansson J., Chen I.M.A., Smirnova T., Nosrat B., Markowitz V.M., Kyrpides N.C. 2012. The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Research 40: D571–D579.

Paszkowski J., Grossniklaus U. 2011. Selected aspects of transgenerational epigenetic inheritance and resetting in plants. Current Opinion in Plant Biology 14: 195–203

Pilate G., Guiney E., Holt K., Petit-Conil M., Lapierre C., Leple J.C., Pollet B., Mila I., Webster E.A., Marstorp H.G., Hopkins D.W., Jouanin L., Boerjan W., Schuch W., Cornu D., Halpin C. 2002. Field and pulping performances of transgenic trees with altered lignification. Nature Biotechnology 20: 607–612.

Plett J.M., Wilkins O., Campbell M.M., Ralph S.G., Regan S. 2010. Endogenous overexpression of *Populus* MYB186 increases trichome density, improves insect pest resistance, and impacts plant growth. Plant Journal 64: 419–432.

Putri S.P., Nakayama Y., Matsuda F., Uchikata T., Kobayashi S., Matsubara A., Fukusaki E. 2013. Current metabolomics: practical applications. Journal of Bioscience and Bioengineering 115: 579–589.

Quesada T., Gopal V., Cumbie W.P., Eckert A.J., Wegrzyn J.L., Neale D.B., Goldfarb B., Huber D.A., Casella G., Davis J.M. 2010. Association mapping of quantitative disease resistance in a natural population of Loblolly pine (*Pinus taeda* L.). Genetics 186: 677–686.

Rabilloud T., Lelong C. 2011. Two-dimensional gel electrophoresis in proteomics: a tutorial. Journal of Proteomics 74: 1829–1841.

Ralph S., Oddy C., Cooper D., Yueh H., Jancsik S., Kolosova N., Philippe R.N., Aeschliman D., White R., Huber D., Ritland C.E., Benoit F., Rigby T., Nantel A., Butterfield Y.S., Kirkpatrick R., Chun E., Liu J., Palmquist D., Wynhoven B., Stott J., Yang G., Barber S., Holt R.A., Siddiqui A., Jones S.J., Marra M.A., Ellis B.E., Douglas C.J., Ritland K., Bohlmann J. 2006. Genomics of hybrid poplar (*Populus trichocarpa* x *deltoides*) interacting with forest tent caterpillars (*Malacosoma disstria*): normalized and full-length cDNA libraries, expressed sequence tags, and a cDNA microarray for the study of insect-induced defences in poplar. Molecular Ecology 15: 1275–1297.

Resende M.D., Resende M.F.Jr., Sansaloni C.P., Petroli C.D., Missiaggia A.A., Aguiar A.M.,

Abad J.M., Takahashi E.K., Rosado A.M., Faria D.A., Pappas G.J., Jr., Kilian A., Grattapaglia D. 2012. Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytologist 194: 116–128.

Resende M.F.Jr., Munoz P., Resende M.D., Garrick D.J., Fernando R.L., Davis J.M., Jokela E.J., Martin T.A., Peter G.F., Kirst M. 2012. Accuracy of genomic selection methods in a standard data set of Loblolly pine (*Pinus taeda* L.). Genetics 190: 1503–1510.

Riesenfeld C.S., Schloss P.D., Handelsman J. 2004. Metagenomics: Genomic analysis of microbial communities. Annual Review of Genetetics 38: 525–552.

Rohde A., Prinsen E., De R.R., Engler G., van M.M., Boerjan W. 2002. PtABI3 impinges on the growth and differentiation of embryonic leaves during bud set in poplar. Plant Cell 14: 1885–1901.

Ruttink T., Arend M., Morreel K., Storme V., Rombauts S., Fromm J., Bhalerao R.P., Boerjan W., Rohde A. 2007. A molecular timetable for apical bud formation and dormancy induction in poplar. Plant Cell 19: 2370–2390.

Saetzler K., Sonnenschein C., Soto A.M. 2011. Systems biology beyond networks: Generating order from disorder through self-organization. Seminars of Cancer Biology 21: 165–174.

Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350–1354.

Sanger F., Nicklen S., Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74: 5463–5467.

Schena M., Shalon D., Davis R.W., Brown P.O. 1995. Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. Science 270: 467–470.

Schmitz R.J., Schultz M.D., Lewsey M.G., O'Malley R.C., Urich M.A., Libiger O., Schork N.J., Ecker J.R. 2011. Transgenerational epigenetic instability is a source of novel methylation variants. Science 334: 369–373.

Schloss P.D., Handelsman J. 2005. Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. Genome Biology 6: 229.

Simon C., Daniel R. 2011. Metagenomic analyses: past and future trends. Applied and Environmental Microbiology 77: 1153–1161.

Singh J., Behal A., Singla N., Joshi A., Birbian N., Singh S., Bali V., Batra N. 2009. Metagenomics: Concept, methodology, ecological inference and recent advances. Biotechnology Journal 4: 480–494.

Tuskan G.A., Difazio S., Jansson S., Bohlmann J., Grigoriev I., Hellsten U., Putnam N., Ralph S., Rombauts S., Salamov A., Schein J., Sterck L., Aerts A., Bhalerao R.R., Bhalerao R.P., Blaudez D., Boerjan W., Brun A., Brunner A., Busov V., Campbell M., Carlson J., Chalot M., Chapman J., Chen G.L., Cooper D., Coutinho P.M., Couturier J., Covert S., Cronk Q., Cunningham R., Davis J., Degroeve S., Dejardin A., Depamphilis C., Detter J., Dirks B., Dubchak I., Duplessis S., Ehlting J., Ellis B., Gendler K., Goodstein D., Gribskov M., Grimwood J., Groover A., Gunter L., Hamberger B., Heinze B., Helariutta Y., Henrissat B., Holligan D., Holt R., Huang W., Islam-Faridi N., Jones S., Jones-Rhoades M., Jorgensen R., Joshi C., Kangasjarvi J., Karlsson J., Kelleher C., Kirkpatrick R., Kirst M., Kohler A., Kalluri U., Larimer F., Leebens-Mack J., Leple J.C., Locascio P., Lou Y., Lucas S., Martin F., Montanini B., Napoli C., Nelson D.R., Nelson C., Nieminen K., Nilsson O., Pereda V., Peter G., Philippe R., Pilate G., Poliakov A., Razumovskaya J., Richardson P., Rinaldi C., Ritland K., Rouze P., Ryaboy D., Schmutz J., Schrader J., Segerman B., Shin H., Siddiqui A., Sterky F., Terry A., Tsai C.J., Uberbacher E., Unneberg P., Vahala J., Wall K., Wessler S., Yang G., Yin T., Douglas C., Marra M., Sandberg G., Van de P.Y., Rokhsar D. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313: 1596–1604.

Vanderschuren H., Lentz E., Zainuddin I., Gruissem W. 2013. Proteomics of model and crop plant species: Status, current limitations and strategic advances for crop improvement. Journal of Proteomics. doi.org/10.1016/j.jprot.2013.05.036

Varshney R.K., Graner A., Sorrells M.E. 2005. Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23: 48–55.

Velasco R., Zharkikh A., Affourtit J., Dhingra A., Cestaro A., Kalyanaraman A., Fontana P., Bhatnagar S.K., Troggio M., Pruss D., Salvi S., Pindo M., Baldi P., Castelletti S., Cavaiuolo M., Coppola G., Costa F., Cova V., Dal Ri A., Goremykin V., Komjanc M., Longhi S., Magnago P., Malacarne G., Malnoy M., Micheletti D., Moretto M., Perazzolli M., Si-Ammour A., Vezzulli S., Zini E., Eldredge G., Fitzgerald L.M., Gutin N., Lanchbury J., Macalma T., Mitchell J.T., Reid J., Wardell B., Kodira C., Chen Z., Desany B., Niazi F., Palmer M., Koepke T., Jiwan D., Schaeffer S., Krishnan V., Wu C., Chu V.T., King S.T., Vick J., Tao Q., Mraz A., Stormo A., Stormo K., Bogden R., Ederle D., Stella A., Vecchietti A., Kater M.M., Masiero S., Lasserre P., Lespinasse Y., Allan A.C., Bus V., Chagne D., Crowhurst R.N., Gleave A.P., Lavezzo E., Fawcett J.A., Proost S., Rouze P., Sterck L., Toppo S., Lazzari B., Hellens R.P., Durel C.E., Gutin A., Bumgarner R.E., Gardiner S.E., Skolnick M., Egholm M., Van de Peer Y., Salamini F., Viola R. 2010. The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42: 833–839.

Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., Smith H.O., Yandell M., Evans C.A., Holt R.A., Gocayne J.D., Amanatides P., Ballew R.M., Huson D.H., Wortman J.R., Zhang Q., Kodira C.D., Zheng X.Q.H., Chen L., Skupski M., Subramanian G., Thomas P.D., Zhang J.H., Miklos G.L.G., Nelson C., Broder S., Clark A.G., Nadeau C., McKusick V.A., Zinder N., Levine A.J., Roberts R.J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z.M., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A.E., Gan W., Ge W.M., Gong F.C., Gu Z.P., Guan P., Heiman T.J., Higgins M.E., Ji R.R., Ke Z.X., Ketchum K.A., Lai Z.W., Lei Y.D., Li Z.Y., Li J.Y., Liang Y., Lin X.Y., Lu F., Merkulov G.V., Milshina N., Moore H.M., Naik A.K., Narayan V.A., Neelam B., Nusskern D., Rusch D.B., Salzberg S., Shao W., Shue B.X., Sun J.T., Wang Z.Y., Wang A.H., Wang X., Wang J., Wei M.H., Wides R., Xiao C.L., Yan C.H., Yao A., Ye J., Zhan M., Zhang W.Q., Zhang H.Y., Zhao Q., Zheng L.S., Zhong F., Zhong W.Y., Zhu S.P.C., Zhao S.Y., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H.J., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M.L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferriera S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland T., Ibegwam C., Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y.H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N.N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J.F., Guigo R., Campbell M.J., Sjolander K.V., Karlak B., Kejariwal A., Mi H.Y., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooseph S., Allen D., Basu A., Baxendale J., Blick L., Caminha M., Carnes-Stine J., Caulk P., Chiang Y.H., Coyne M., Dahlke C., Mays A.D., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J.,

Kagan L., Kraft C., Levitsky A., Lewis M., Liu X.J., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., Nodell M., Pan S., Peck J., Peterson M., Rowe W., Sanders R., Scott J., Simpson M., Smith T., Sprague A., Stockwell T., Turner R., Venter E., Wang M., Wen M.Y., Wu D., Wu M., Xia A., Zandieh A., Zhu X.H. 2001. The sequence of the human genome. Science 291: 1304–1351.

Venter J.C., Remington K., Heidelberg J.F., Halpern A.L., Rusch D., Eisen J.A., Wu D.Y., Paulsen I., Nelson K.E., Nelson W., Fouts D.E., Levy S., Knap A.H., Lomas M.W., Nealson K., White O., Peterson J., Hoffman J., Parsons R., Baden-Tillson H., Pfannkoch C., Rogers Y.H., Smith H.O. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science 304: 66–74.

Verde I., Abbott A.G., Scalabrin S., Jung S., Shu S., Marroni F., Zhebentyayeva T., Dettori M.T., Grimwood J., Cattonaro F., Zuccolo A., Rossini L., Jenkins J., Vendramin E., Meisel L.A., Decroocq V., Sosinski B., Prochnik S., Mitros T., Policriti A., Cipriani G., Dondini L., Ficklin S., Goodstein D.M., Xuan P., Del Fabbro C., Aramini V., Copetti D., Gonzalez S., Horner D.S., Falchi R., Lucas S., Mica E., Maldonado J., Lazzari B., Bielenberg D., Pirona R., Miculan M., Barakat A., Testolin R., Stella A., Tartarini S., Tonutti P., Arus P., Orellana A., Wells C., Main D., Vizzotto G., Silva H., Salamini F., Schmutz J., Morgante M., Rokhsar D.S. 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nature Genetics 45: 487–494.

Wang Z., Gerstein M., Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Review in Genetics 10: 57–63.

Weckwerth W. 2011. Green systems biology – From single genomes, proteomes and metabolomes to ecosystems research and biotechnology. Journal of Proteomics 75: 284–305.

Wegrzyn J.L., Eckert A.J., Choi M., Lee J.M., Stanton B.J., Sykes R., Davis M.F., Tsai C.J., Neale D.B. 2010. Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. New Phytologist 188: 515–532.

Wilm M., Shevchenko A., Houthaeve T., Breit S., Schweigerer L., Fotsis T., Mann M. 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379: 466–469.

Wilmes P., Bond P.L. 2006. Metaproteomics: studying functional gene expression in microbial ecosystems. Trends in Microbiology 14: 92–97.

Wullschleger S.D., Jansson S., Taylor G. 2002. Genomics and forest biology: *Populus* emerges as the perennial favorite. Plant Cell 14: 2651–2655.

Yakovlev I.A., Fossdal C.G., Johnsen O. 2010. MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. New Phytologist 187: 1154–1169.

Yates J.R., Ruse C.I., Nakorchevsky A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. Annual Review of Biomedical Engineering 11: 49–79.

Yordanov Y., Regan S., Busov V. 2010. Members of the Lateral Organ Boundaries Domain (LBD) transcription factors family are involved in regulation of secondary growth in *Populus*. Plant Cell 22: 3662–3677.

Zawaski C., Kadmiel M., Ma C., Gai Y., Jiang X., Strauss S.H., Busov V.B. 2011a. SHORT INTERNODES-like genes regulate shoot growth and xylem proliferation in *Populus*. New Phytologist 191: 678–691.

Zawaski C., Kadmiel M., Pickens J., Ma C., Strauss S., Busov V. 2011b. Repression of gibberellin biosynthesis or signaling produces striking alterations in poplar growth, morphology, and flowering. Planta 234: 1295–1298.

Zhang F., Maeder M.L., Unger-Wallace E., Hoshaw J.P., Reyon D., Christian M., Li X.H., Pierick C.J., Dobbs D., Peterson T., Joung J.K., Voytas D.F. 2010. High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proceedings of the National Academy of Sciences of the United States of America 107: 12028–12033.