DENDROBIOLOGY
2012, vol. 68, 113–123

*Aleksandra M. Staszak, Tomasz A. Pawłowski**

Forest tree research in post genomic era. Introduction to systems biology of broadleaves

Received: 16 February 2012; Accepted: 28 June 2012

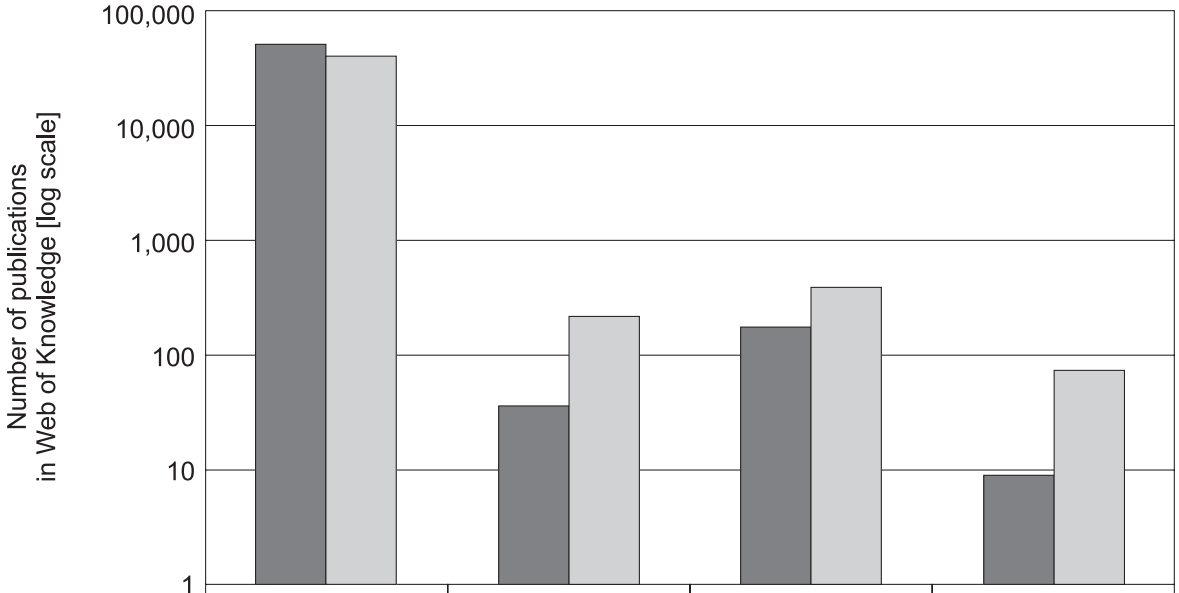
Abstract: Trees are long living organisms, rarely used in molecular experiments because of large size of the genome and long time of reproduction cycle. Sequencing data from *Populus trichocarpa* genome allowed for the development of research on the processes associated with tree biology such as secondary wood formation, long-term perennial growth, seasonal changes, biotic interactions, evolution etc. Reference data enable the investigation of non-model trees such as *Quercus* or *Fagus*, having ecological and economic significance. During projects scientists use genomic, transcriptomic, proteomic and metabolomic approaches which contribute to better understanding of the physiological processes regulating tree biology. Data collected from these multiple studies need to be integrated. The integration of data is the subject of the newly established field of science called systems biology. This review presents progress in tree research after finishing the sequencing project of *Populus*. It concentrates on modern trends in 'omics' and systems biology study of temperate broadleave trees during the last 10 years of studies.
Abbreviations: DE – dimensional electrophoresis, CE – capillary electrophoresis, DIGE – differential 2DE, EST – expressed sequence tag, FT – ICR – MS – Fourier transform ion cyclotron resonance mass spectrometry, GC – gaschromatography, LC – liquid chromatography, MALDI-TOF – matrix-assisted laser desorption/ionization time of flight, MPSS – massive parallel signature sequencing, MS – mass spectrometry, MW – mass weight, NCBI EGP – National Center for Biotechnology Information Entrez Genome Project, NMR – nuclear magnetic resonance, HPLC – highperformance liquid chromatography, SNP – single nucleotide polymorphisms, SRS – short reads sequences, SSH – suppression subtractive hybridization, TFA – transcriptome fingerprinting analysis, qRT-PCR quantitative Real Time PCR, QTL – quantitative trait loci

Additional key words: genomics, proteomics, metabolomics, transcriptomics, woody plants

Address: Aleksandra M. Staszak, Tomasz A. Pawłowski Institute of Dendrology Polish Academy of Sciences, Kórnik, Poland, e-mail: tapawlow@man.poznan.pl
*corresponding author

Introduction

Trees are an important group of plants with environmental and economic significance. Because of the long life cycle and big size of the genome trees are not often used in basic research carried out on molecular level. They were used as a model plant by not so many experimental groups but this situation has been changing after sequencing of genome of the first tree species *Populus trichocarpa* in 2006 (Tuskan et al. 2006; see Fig. 1). Completion of this project has contributed to better understanding of the structure of the genome and to the development of a new field of research. On the basis of this 'genomic era' the 'omic' sciences came into being. Nowadays in the 'post genomic era' scientists in tree research commonly use 'omic' approaches. Proteomics develops faster than other sciences like trancriptomics or metabolomics.



	gen* tree*	transcriptom* tree*	proteom* tree*	metabolom* tree*
before Oct 2006	51,289	36	179	9
after Oct 2006	40,300	217	390	75

Fig. 1. Number of 'omic' publications published before and after *Populus trichocarpa* genome sequencing concerning tree science

Bioinformatics enables collection and compilation of huge volumes of data coming from these investigations. Altogether, 'omic' sciences offer a possibility of describing processes not at one but at many different levels, building an overall model of organisms functionality. Such understanding is the basis of a new field of study, the systems biology, which in a holistic way combines data from different areas of research.

Herein we present the progress in tree research after completion of the sequencing project of a model tree *Populus*. The review concentrates on modern trends in genomics, transcriptomics, proteomics and metabolomics of temperate broadleave trees during the last decade of studies. Table 1 presents research papers reviewed in this work.

The genome of forest trees: what do we know?

Arabidopsis thaliana was the first known plant genome, published in 2000 (Arabidopsis Genome Initiative). Six years later, the first tree species genome of poplar (*Populus trichocarpa*) was published. Sequencing is a tool for the genome structure recognition being the basis for the next generation investigation of the role and function of genes. In tree research, *Populus* has been chosen because of the rapid growth rate, relatively small size of the genome (~485 Mb) and ecological and economic importance. Such a model allows the scientists to study many processes associated with tree biology e.g. dormancy, secondary wood formation, long-term perennial growth, seasonal changes, flowering, reproduction, biotic interactions, evolution of adaptive traits and speciation. *Populus* can be genetically transformed. Interspecific *Populus* hybrids are phenotypically diverse. The change in phenotype and the data from the sequencing project help in gene mapping and establishment of their functions. The data can also be used in practice, in breeding economically important hybrids. Because of the fast growth rate, some hybrids could be bred on a large scale, as a source of renewable energy (Bradshaw et al. 2000, Peńa and Séguin 2001, Taylor 2002, Tuskan et al. 2003, Tuskan et al. 2006, Jansson and Douglas 2007).

At present, apart from *Populus* no other forest species of trees or shrubs can be found in the completed large-scale sequencing projects. Similarly, no forest species can be found in the large-scale sequencing projects currently in progress. The trees for which only genetic maps have been established are *Corylus avelana*, *Quercus robur* and *Salix viminalis* (National Center for Biotechnology Information Entrez Genome Project, NCBI EGP).

The database of the expressed sequence tags (ESTs) contains much more information concerning genomes of forest tree species. Properties like: lower cost, shorter time to generate EST data (versus traditional sequencing) and the fact that the resulting data may be useful in future research encourage the researchers to use this technique (Ohlrogge and Ben-

Table 1. Forest tree research papers reviewed in this work.

Aim of study	Species	Part of plant	Techniques	Remarks	References
Development					
bud burst	*Quercus petraea*	buds	QTL, EST, qRT-PCR, SSH, microarrays	TF DAG2 is involved in buds dormancy release	Derory et al. (2006)
cell wall lignifications	*Populus tremula* × *alba* and two transgenic lines	cambium, leaf tissue cultures	GC-MS, HPLC	Authors defined metabolic compounds of lignin that links gene expression and phenotype.	Robinson et al. (2005)
cellulose synthesis	*P. deltoides* × *trichocarpa*	xylem	2DE, LC-MS /MS	Two types of cellulose synthase complexes (CSC) participate in secondary wall formation.	Song et al. (2010)
leaves sesonal changes	*P. tremula* × *P. tremuloides*	leaves	microarrays, GC-TOF-MS	Changes in transcriptome and metabolome profiles of leaves during long and short days are differently regulated.	Hoffman et al. (2010)
root development	*P. trichocarpa* × *P. deltoids*	root	EST	The authors have created root EST databases for the investigation of changes in aquaporins and transporter transcripts.	Kohler et al. (2003)
wood formation	*P. tremula L.* × *P. tremuloides* Michx. and *P. trichocarpa* 'Trichobel.'	leaves	EST	The authors created a cambium cDNA library from *P. tremula* x tremuloides, and the developing xylem cDNA library from *P. trichocarpa*.	Sterky et al. (1998)
xylem development	*P. grandidentata* × *alba, P. tremula* × *alba, Pseudotsuga menziesii, Pinus radiata*	xylem, cambium	GC/MS	Revision of the procedure of lignin monomers extraction.	Robinson (2009)
Stress response					
drought	*P. trichocharpa, P. trichocarpa* × *P. deltoides*	bud, root, inflorescence, leaves	2DE, LC-MS/MS, MALDI-TOF-MS	Studies involved 'subproteomes' interrogation of e.g. cell wall, plasma membrane, vacuolar membrane, ER, golgi apparatus, mitochondrion and chloroplasts.	Plemion et al. (2006)
drought	*P. tremula* L.× *P. alba* L. (P. × canescens (Aiton) Smith)	leaves, cambium	2DE, MS/MS	Drought stress response in cambium is faster than in leaves. Changes in the cambial proteome after rewatering disappeared whereas in leaves many proteins appeared to be differentially regulated only during the recovery from drought.	Durand et al. (2011)
drought	*Quercus robur* L.	leaves	analysis of carbohydrate accumulation, 2D DIGE	Oak initially adapted its metabolism in order to maintain the full molecular functionality. However prolonged drought exposure overwhelmed the adaptive mechanisms.	Sergeant et al. (2011)
gene response to Zn	*P.* × *euramericana* (Dode)	leaves	microarrays	The characterization of genes that are activated in poplar response to Zn allows for the selection of clones for remediation technologies.	Di Baccio et al. (2011)
ozone exposure	*P. trichocarpa* Torr. & Gray, *P. deltoides* Bart	leaves	microarray, QTL	QTL mapping identified regions involved in expression to ozone which were found to co-localise to QTL for necrotic damage, providing encouraging evidence for their importance in governing this trait.	Street et al. (2011)
ozone exposure	*Fagus sylvatica* L.	leaves	2D DIGE	Under ozone exposure, abundance of proteins related to the Calvin cycle and photosynthetic electron transport chain were decreased whereas the abundance of proteins regarding the carbon metabolism/catabolism were increased.	Kerner et al. (2011)

Aim of study	Species	Part of plant	Techniques	Remarks	References
ozone exposure	P. tremula L. × *P. alba* L. (*Populus* x *canescens* (Aiton) Smith	chloroplast membrane	DIGE	PS and ATPase subunits decrease in abundance could be the result of oxidative processes on chloroplast proteins but could also be a way to down-regulate photochemical reactions in response to an inhibition in Calvin cycle activity.	Bohler et al. (2011)
resistance to insects	*Fraxinus pennsylvanica*, *F. americana*, *F. nigra*, *F. quadrangulata*, *F. mandshurica*	phloem	454 pyrosequencing	The data provide an invaluable resource for understanding the genetic make-up of ash phloem, the target tissue of *Agrilus planipennis*.	Bai et al. (2011)
resistance to insects	*Fraxinus* spp.	phloem	DIGE	Genes involved in constitutive resistance to the Emerald Ash Borer: PR-10 protein, an aspartic protease, PCBER, and ascorbate peroxidase.	Whitehill et al. (2011)
salt sensitivity	*P. x canescens*	leaves	histology, x-ray microanalysis qRT-PCR	Authors analyzed expression patterns of transport proteins associated with ion concentrations, uptake and transport.	Escalante-Pérez et al. (2009)
salt sensitivity	*P. euphratica, P.* × *canescens*	leaves	FT-ICR-MS, microarray analysis, qRT-PCR,	The evolutionary adaptation of *P. euphratica* to saline environments is apparently linked with higher energy requirement of cellular metabolism and a loss of transcriptional regulation.	Janz et al. (2010)
salt sensitivity	*P. euphratica*	living tissues, callus	EST	This study represents the deepest transcriptomic and gene-annotation analysis of *P. euphratica* to date.	Qiu et al. (2011)
salt stress	*P. alba* L.	leaves	microarrays, qRT-PCR	71 genes were functionally related to carbohydrate metabolism, energy metabolism and photosynthesis.	Beritognolo et al. (2011)
salt stress	*P. euphratica*	leaves, roots	qRT-PCR, GC-MS	Adaptation to saline condition is connected with gene differential expression.	Brosché et al. (2006)
water deficit	*P. alba* L.	cambium	Microarray analysis, qRT-PCR	The water deficit resulted in changes in gene expression of protein metabolism, cell wall metabolism, stress response, transporters and transcriptional regulation.	Berta et al. (2010)
tissue regeneration after bark girdling	*P. tomentosa* (Carrière)	bark	microarrays, qRT PCR	Differentiating xylem cells acquire regenerative competence through epigenetic regulation and cell cycle re-entry. The xylem developmental program was blocked, whereas the phloem or cambium program was activated. Phytohormones play important roles in vascular tissue regeneration	Zhang et al. (2011)
Others					
genetic variability	*Q. ilex* subsp. *Ballota*	acorn	2DE, MS/MS	The authors created protein maps of acorns and characterized natural biodiversity in 10 populations.	Valero Galván et al. (2011)
pollinosis	*Birch verrucosa*	pollen	1DE, 2DE, nano-LC-MS/MS	Molecular characterization of pollen extracts is relevant for standardization and development of new reagents for specific immunotherapy.	Erler et al. (2011)
pollinosis	*Birch* sp.	pollen	2DE, LC-MS	Differences in Bet v 1 composition has no effect on changes in the allergenicity.	Schenk et al. (2011)

ning 2000). The ESTs database (http://www.ncbi.nlm.nih.gov/dbEST) is a powerful tool for collecting coding content and expression patterns for different tissues, environments and species (Sterky et al. 2004). The first use of ESTs data was to identify the genes involved in plant metabolic pathways. As at December 1, 2011 71,276,166 sequences were stored in dbEST. Databases cover information about ESTs of the following tree species: *Populus*, *Picea*, *Pinus*, *Quercus*, *Salix*, *Fagus*, *Taxus*, *Betula*, *Pseudotsuga* and *Alnus* (Ueno et. al. 2010, Rigault et al. 2011). ESTs are used in different kinds of research e. g. in transcriptome and metabolome profiling.

In recent times, genomics relies on the combination of traditional genetic methods and the tools used in the 'omic' studies. For example, Derory et al. (2006) combined Quantitative Trait Loci (QTL) with Expressing Sequence Tag (EST), Quantitative Real Time PCR (qRT-PCR) and Suppression Subtractive Hybridization (SSH) in studies on genes differentially expressed between the quiescent and active stage of oak bud development. This research confirmed the usefulness of combination of such methods in the identification of relevant candidate genes. Gailinga et al. (2009) used single nucleotide polymorphisms (SNPs) with functional genomics protocol to assess adaptive genetic variation in oak.

Transcriptomics

The transcriptomic approach is based on the analysis of gene expression in certain locations (e. g. in leaves, roots, groups of cells) and in time. The transcriptomes of organisms change dynamically depending on the environmental conditions (e. g. biotic and abiotic stresses, Brosché et al. 2006), stage of life cycle, or seasonal shifts. Research focuses mostly on model organisms, such as *Arabidopsis thaliana*, or as for trees, on *Populus trichocarpa*. In tree research, different topics concerning the *Populus* genus have been studied. The transcriptomic approach was used e. g. in research on wood-formation (Sterky et al. in 1998), root growth and water-stress (Kohler et al. 2003), seasonal changes in leaves (Sjödin et. al. 2006, Hoffman et al. 2010), salt sensitivity (Escalante-Pérez et al. 2009, Janz et al. 2010, Qiu et al. 2011), water deficit (Plemion et al. 2006, Berta et al. 2009), and response to ozone exposure (Street et al. 2011). Sjödin et. al. (2006) created a database named UPSC-BASE containing transcriptomic data from *P. trichocarpa*.

The next generation of new sequencing methods will provide a possibility to develop research on non-model trees species. The analysis of results collected from the 'omic' research is limited by insufficient number of references in the databases.

Nowadays, transcriptomic analysis is based mainly on microarrays and qRT-PCR, which allow to study the expression of known genes. Microarrays are used in measuring the level of large numbers of genes simultaneously. qRT-PCR is used for gene expression comparisons on a small scale. The EST library is used on a larger scale in studies carried out on gene expression profiling (Sterky et al. 2004). A high cost of those protocols has led to development of other methods like Massive Parallel Signature Sequencing (MPSS) which enables analysis of the gene expression in a sample by counting the number of individual mRNA molecules produced by each gene (Brenner et al. 2000). Using next generation sequencing technique (NGS; Stapley et al. 2010) makes it possible to produce even a million of sequences in one run. The Roche 454 FLX Titanium system, Illumina's Genome Analyser (Solexa), ABI's SOLiD platforms, HeliScope, Ion Torrent, PacBio and Stright are used nowadays in transcriptome studies (Glenn 2011). The high costs of all the listed protocols have driven researchers to look for new protocols less expensive, laborous and time consuming. For example, Transcriptome Fingerprinting Analysis (TFA) allows detection of gene expression patterns in studies of picoeukaryotic marine microbial communities. The TFA has emerged as a tool for indication of changes in the samples and their pre-selection before using more powerful, time-intensive and costly methods (Coll-Lladó et al. 2011). The Roche 454 FLX Titanium system was used in the research of non-model organisms *Melitaea cinxia* and Nymphalidae (Vera et al. 2008). The Short-Read Sequences (SRS) protocol was used for the investigation of *Pachycladon enysii* (Collins et al. 2008). The Collins group (2008) has mapped *Pachycladon* orthologues to specific *A. thaliana* loci, in order to find putative duplicate of *Pachycladon* genes. The use of SRS to compare the sequences of species without a close reference is difficult but possible. Improvement of the algorithms is required (Surget-Groba et al. 2010). The study of trees, also non-modelspecies, can be based on the next generation methods or on newly generated techniques which will be discovered in the future.

Proteomics

The proteomic approach is the fastest developing one from all the "omic" sciences. Proteomics complements analysis of the transcriptome and the metabolome. It is an essential source of information about biological systems because it generates knowledge about the concentrations, interactions, functions, and catalytic activities of proteins, which are the major structural and functional determinants of cells (Baginsky 2009).

The proteomics of trees is a fast developing area of research, yet we are far from expectation of full understanding of the role of proteins in tree biology. Until 2011, the proteomic approach was used for re-

search on tree species such as *Populus*, *Pinus*, *Eucalyptus*, *Picea*, *Fagus, Quercus*, *Acer, Hevea* and *Cunninghamia*. The results of those studies were widely discussed by Abril et al. (2011). In this review, only the data which are not mentioned in Abril's review will be presented. Miernyk and Hajduch (2011) have reviewed the publications concerning proteomics of seeds, including the seeds of trees. However, they have concentrated mostly on the storage proteins.

The tree scientists' attention is focused on studying the proteome profiles of pollen (Erler et al. 2011, Schenk et al. 2011), leaves, roots or seeds (Pawłowski and Kalinowski 2003; Szczotka et al. 2003, Pawłowski 2007, 2009, 2010). Pawłowski (2007, 2009) identified functional proteins associated with the Norway maple (*Acer platanoides*) and beech (*Fagus sylvatica*) seeds dormancy breaking. Most of the proteins were under control of abscisic and gibberellic acids, hormones regulating dormancy status (Pawłowski 2010). Furthermore, the proteomic techniques are used to investigate processes associated with ozone exposure (Kerner et al. 2011), xylem tissues forming (Song et al. 2010) or stress influence (Durand et al. 2011).

The majority of research was done using electrophoresis: one dimensional (1DE), two dimensional (2DE) (reviewed by Jorrín-Novo et al. 2009) or differential 2DE (DIGE) (Tonge et al. 2001). More and more attention of researchers is focused on quantitative MS technologies (Oeljeklaus et al. 2009). Different techniques of MS can be used as powerful tools for research carried on plant material. Tandem mass spectrometry (MS/MS) is used e.g. in phospoproteome analysis (Palumbo et al. 2011).

Proteomic tools have been divided into classical (based on gel) and second generation (gel and label free) magnifying capabilities of protein coverage. Multidimensional chromatography protein identification (MudPIT) makes use of isotope labeling which allows for investigation of post-translational modifications, high-throughput protein identification and investigation of quantitative differences in protein expression (Cańas et al. 2007).

Metabolomics

The metabolome is understood as a complete set of small molecules (i.e. metabolites) which participate in, or are products of, metabolic reactions within an organism or tissue. The metabolomic profile reflects changes in the plant (usually in certain places such as the root, leaf, flower, seed etc.) which might be e. g. in different stages of the life cycle or stress conditions. The advantage of metabolomics is that it can be applied to non-model plants without a need for the genome information, however genome sequences are sometimes used to predict the occurrence of metabolites. The design of such studies generally includes plant cultivation, sampling, extraction, derivatization, separation and quantification, data matrix conversion, data mining, and bioscience feedback which can involve a lot of experimental errors (Fukusaki and Kobayashi 2005).

Metabolomic data are collected in numerous studies together with transcriptomic and proteomic data to show a more global pattern of changes. Because of the lack of standard methods used in this type of research, nowadays metabolomics is rarely used alone (Ward et al. 2007, Wienkoop et al. 2008). Robinson et al. (2005) have examined the potential of metabolite profiling as a selection tool for genotype discrimination in *Populus*.

Different techniques are used to study the metabolomic profiles of plants. The scientists use chromatography (gas chromatography GC for the analysis of small molecules MW < 1000, or alternatively, high performance liquid chromatography HPLC for the analysis of large or labile molecules), mass spectrometry (coupled with gas chromatography GC-MS or capillary electrophoresis CE-MS for the analysis of hydrophilic small molecules, Fourier transform ion cyclotron resonance FT-ICR-MS for all purposes, nuclear magnetic resonance NMR spectroscopy (used as an alternative to chromatography/mass spectrometry due to its main advantage of having a non-destructive effect on the sample) and vibrational spectroscopic technique (Dunn et al. 2005, Fukusaki and Kobayashi 2005, Robinson 2009, Janz et al. 2010, Ward et al. 2010). Regardless of the chosen technique of metabolomic study, eventually all paths lead to identifying (and quantifying) the key metabolites. Given the chemical diversity of metabolomes, metabolite identification is intrinsically difficult (Wishart 2011). Recently a lot of attention is given to NMR spectroscopy (Kim et al. 2011). 1D NMR is used in the classification of similar groups of samples while 2D NMR is used to characterize unidentified compounds from the 1D protocol. LC-NMR allows for chemical characterization of samples and its advanced version, LC-SPE-NMR (liquid chromatography – solid phase extraction – nuclear magnetic resonance), offers the possibility of examination of alcoholic extracts and identification of flavanolglycosides and cardenolides (Ward et al. 2007).

Systems biology

The results of studies are like pieces of a jigsaw puzzle, which together create a picture of the global understanding of what happens in the living organisms. To understand biology at the systems level, we must examine the structure and dynamics of the cellular and organism functions, rather than the characteristics of isolated parts of a cell or an organism (Kitano 2002).

This is the fundamental concept of the systems biology. Such understanding requires amalgamation of data from all kinds of studies, not only proteomic or metabolomic but also ecological, physiological etc. Sjödin (2007) claims that understanding of the complexity of biological processes is possible only in the case of integration of knowledge from different fields of biological science. Kitano (2002) divides the systems biology research into four key directional categories: structure, dynamics, design and control methods needed to avoid erros. The integration has to occur on each level starting from the basic genome data through transcriptome, proteome and metabolome to the phenotype. Such understanding of all processes taking place in different organs and plants can give the scientists answers to the fundamental questions concerning the functioning of organisms.

In the field of tree research, most of the integrative data (coming from a limited number of studies) concern *Populus* species. Broad-range metabolomic and transcriptomic studies carried by Sjödin (2007) on development and autumn senescence of *Populus* leaves suggest that processes are much more complex than we thought. Hoffman et al. (2010) have chosed the aspen hybrid (*P. tremula* x *P. tremuloides*) to study the impact of seasonal photoperiod, environmental signal that affects many physiological changes in plants e.g. timing of the winter dormancy. They have integrated transcriptomic and metabolomic data and pointed out that 16% of the genes were diurnally regulated. Several of these genes were involved in circadian-associated processes, including photosynthesis and primary and secondary metabolism. Metabolites were mostly involved in carbon metabolism. Direct linking of the transcript changes with the changes in metabolite profiles was very difficult. Changes in the metabolome may occur later than the corresponding transcriptomal changes. Further investigation is required to elucidate the mechanisms involved in plants' adaptation to new photoperiods at transcript and metabolite levels more comprehensively. Differences in transcriptome and metabolome in *P. euphratica* (salt tolerant) and *P. x canescens* (salt sensitive) were the focus of Janz et al. (2010) studies. The evolutionary adaptation of *P. euphratica* to saline environments was apparently linked with higher energy requirement of cellular metabolism and a loss of transcriptional regulation.

Discussion

Twelve years ago the first plant genome of *Arabidopsis thaliana* was released. Many researchers expected that completion of sequencing projects would give answers to many questions which they had asked before sequencing started. However, the projects deliver a lot of data which do not offer easy answers but rather give rise to new questions and hypotheses about genome functionality. The sequencing program has already proven to be a milestone in molecular biology. It has opened a new post-genomic era, that is linked with the emergence of a new research field in biology. Nowadays 'omics' sciences are strictly related to the genome of organisms. Genomics, transcriptomics, proteomics and metabolomics are the main directions of research. Currently, all the 'omic' sciences develop rapidly and dynamically. Following the commencement of sequencing programs for forest species (the first one being *Populus trichocarpa*), 'omic' sciences have also conquered this area. Researchers discuss about the structure of 'omic' sciences. Plemion et al. (2006) treat proteomics as an active field of genomics, while Remmerie et al. (2001) claim that functional proteomics and proteogenomics stem from from functional genomics and have nowadays become an equal player in the systems biology. Abril et al. (2011) have stated that proteomic is a fundamental discipline in the post-genomic era. At present, 'omic' sciences branch out into specialized subjects e.g. proteomics into modificomics (phosphoproteomics, Kersten et al. 2006, Peck 2006 or glycoproteomics, Hashii et al. 2005, Remmerie et. al 2011), while interactomics (protein-protein interactions) is treated like a separate discipline, not part of proteomics (Lo 2007, Ivanov et al. 2011). Fluxomics is a branch of metabolomics (Kim et al. 2011).

A search through databases could return more and more information about new 'omics' sciences like glycomics (study aimed at comprehensive elucidation and characterization of all the glycoforms like monosaccharides, oligosaccharides, polysaccharides, and their modifications, Gupta et al. 2009). The progress in post-genomic area leads to formation of new 'omics' sciences almost each day.

Nowadays we can take advantage of more and more advanced databases and web tools, which help not only in the planning of experiments but also in the understanding of their results. NCBI created a wide range of databases including PubMed – bibliographic databases and PuBMed Central, GenBank – nucleotide sequence databases, (dbEST, dbSTS, or dbGSS), Molecular Structure Databases, Taxonomy databases and Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation etc (McEntyre and Ostell 2002). The Internet offers access to different databases e.g. PROTICdb in proteomics (Ferry-Dumazet et al. 2005), or UPSC- BASE in transcriptomics (Sjödin et al. 2006). The databases of NMR metabolomic analyses still contain a limited number of records (Kim et al. 2011). In 2001, a database called Babelomics was set up. Babelomics is an integrative platform for the analysis of transcriptomic, proteomic and genomic data with advanced functional profiling (http://babelomics.bioinfo.cipf.es, reviewed by Al-Shahrour et al. 2005, 2007). Babe-

lomics helps the scientific community offering an advanced set of methods for the integrated analysis of genomic data (Medina et al. 2010).

Conclusions and future perspectives

Sequencing of Arabidopsis and subsequently *Populus* led to the development of a new generation of techniques useful for the investigation of other non-model but economically and ecologically important tree species. The understanding of processes associated with tree biology can be more effective, because of parallel development of a new area of research delivering data for genomic resources. New branches of research are developing as modificomics, interactomics etc. Structured data from e.g. crystallography and other area of research (such as physics) can be integrated with data from biological research to give a more comprehensive picture of processes taking place in organisms. On the other hand, the growing resources of information goffering researchers more possibilities to search for, compare and publish the results of their work, give rise to a need for special engines and interpretation software. Such tools should allow for compilation of data from different 'omic' studies and facilitate the building of a universal model of organism functionality. Collaboration of the scientific community seems to be the foundation for the development of the systems biology.

Acknowledgments

The work was supported by the National Science Centre, Poland, grant number 2011/01/B/NZ9/02868.

References

Abril N., Gion J.M., Kerner R., Muller-Starck G., Cerrillo R.M.N., Plomion C., Renaut J., Valledor L., Jorrin-Novo J.V. 2011. Proteomics research on forest trees, the most recalcitrant and orphan plant species. Phytochemistry 72: 1219–1242.

Al-Shahrour F, Minguez P, Vaquerizas JM, Conde L, Dopazo J. 2005. Babelomics: a suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments. Nucleic Acids Research 33: 460–464.

Al-Shahrour F., Arbiza L., Dopazo H., Huerta-Cepas J., Mínguez P., Montaner D., Dopazo J. 2007. From genes to functional classes in the study of biological systems. BMC Bioinformatics 8: Article Number 114.

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815.

Baginsky S. 2009. Plant proteomics: Concepts, applications, and novel strategies for data interpretation. Mass Spectrometry Review 28: 93–120.

Bai X.D., Rivera-Vega L., Mamidala P., Bonello .P, Herms D.A., Mittapalli O. 2011. Transcriptomic signatures of ash (*Fraxinus* spp.) Phloem. PLoS ONE 6: Article Number e16368.

Beritognolo I., Harfouche A., Brilli F., Prosperini G., Gaudet M., Brosché M, Salani F., Kuzminsky E., Auvinen P., Paulin L., Kangasjärvi J., Loreto F., Valentini R., Scarascia-Mugnozza G., Sabatti M. 2011. Comparative study of transcriptional and physiological responses to salinity stress in two contrasting *Populus alba* L. genotypes. Tree Physiology 31: 1335–1355.

Berta M., Giovannelli A., Sebastiani F., Camussi A., Racchi M.L. 2010. Transcriptome changes in the cambial region of poplar (*Populus alba* L.) in response to water deficit. Plant Biology 12: 341–354.

Bohler S., Sergeant K., Hoffmann L., Dizengremel P., Hausman J.F., Renaut J., Jolivet Y. 2011. A difference gel electrophoresis study on thylakoids isolated from Poplar leaves reveals a negative impact of ozone exposure on membrane proteins. Journal of Proteome Research 10: 3003–3011.

Bradshaw H.D., Ceulemans R., Davis J., Stettler R. 2000. Emerging model systems in plant biology: Poplar (*Populus*) as a model forest tree. Journal of Plant Growth Regulation 19: 306–313.

Brenner S., Johnson M., Bridgham J., Golda G., Loyd D.H., Johanson D., Luo S., McCurdy S., Foy M., Ewan M., Roth R., George D., Eletr S., Albrecht G., Vermaas E., Williams S.R., Moon K., Burcham T., Pallas M., DuBridge R.B., Kirchner J., Fearon K., Mao J., Corcoran K. 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nature Biotechnology 18: 630–634.

Brosché M., Vinocur B., Alatalo E.R., Lamminmäki A., Teichmann T., Ottow E.A., Djilianov D., Afif D., Bogeat-Triboulot M.B., Altman A., Polle A., Dreyer E., Rudd S., Paulin L., Auvinen P., Kangasjärvi J. 2005. Gene expression and metabolite profiling of *Populus euphratica* growing in the Negev desert. Genome Biology 6: Article Number R101.

Cañas B., Piñeiro C., Calvo E., López-Ferrer D., Gallardo J.M. 2007. Trends in sample preparation for classical and second generation proteomics. Journal of Chromatography A 1153: 235–258.

Collins L.J., Biggs P.J., Voelckel C., Joly S. 2008. An approach to transcriptome analysis of non-model organisms using short-read sequences. Genome Informatics 21: 3–14.

Coll-Lladó M., Acinas S.G., Pujades C., Pedrós-Alió C. 2011. Transcriptome fingerprinting analysis: an

approach to explore gene expression patterns in marine microbial communities. PLoS ONE 6: Article Number e22950.

Derory J., Léger P., Garcia V., Schaeffer J., Hauser M.T., Salin F., Luschnig C., Plomion C., Glössl J., Kremer A. 2006. Transcriptome analysis of bud burst in sessile oak (*Quercus petraea*). New Phytologist 170: 723–738.

Di Baccio D., Galla G., Bracci T., Andreucci A., Barcaccia G., Tognetti R., Sebastiani L. 2011. Transcriptome analyses of *Populus* × *euramericana* clone I-214 leaves exposed to excess zinc. Tree Physiology 31: 1293–1308.

Dunn W.B., Bailey N.J.C., Johnson H.E. 2005. Measuring the metabolome: current analytical technologies. Analyst 130: 606–625.

Durand T.C., Sergeant K., Renaut J., Planchon S., Hoffmann L., Carpin S., Label P., Morabito D., Hausman J.F. 2011. Poplar under drought: comparison of leaf and cambial proteomic responses. Journal of Proteomics 74: 1396–1410.

Escalante-Pérez M., Lautner S., Nehls U., Selle A., Teuber M., Schnitzler J.P., Teichmann T., Fayyaz P., Hartung W., Polle A., Fromm J., Hedrich R., Ache P. 2009. Salt stress affects xylem differentiation of grey poplar (*Populus* x *canescens*). Planta 229: 299–309.

Erler A., Hawranek T., Kruckemeier L., Asam C., Egger M., Ferreira F., Briza P. 2011. Proteomic profiling of birch (*Betula verrucosa*) pollen extracts from different origins. Proteomics 11: 1486–1498.

Ferry-Dumazet H., Houel G., Montalent P., Moreau L., Langella O., Negroni L., Vincent D., Lalanne C., de Daruvar A., Plomion C., Zivy M., Joets J. 2005. PROTICdb: A web-based application to store, track, query, and compare plant proteome data. Proteomics 5: 2069–2081.

Fukusaki E., Kobayashi A. 2005. Plant metabolomics: potential for practical operation. Journal of Bioscience and Bioengineering 100: 347–354.

Gailing O., Vornam B., Leinemann L., Finkeldey R. 2009. Genetic and genomic approaches to assess adaptive genetic variation in plants: forest trees as a model. Physiologia Plantarum 137: 509–519.

Glenn T.C. 2011. Field guide to next-generation DNA sequencers. Molecular Ecology Resources 11: 759–769.

Gupta G., Surolia A., Sampathkumar S.G. 2010. Lectin microarrays for glycomic analysis. OMICS: a Journal of Integrative Biology14: 419–436.

Hashii N., Kawasaki N., Itoh S., Hyuga M., Kawanishi T., Hayakawa T. 2005. Glycomic/glycoproteomic analysis by liquid chromatography/mass spectrometry: Analysis of glycan structure alternation in cells. Proteomics 5: 4665–4672.

Hoffman D.E., Jonsson P., Bylesjö M., Trygg J., Antti H., Eriksson M.E., Moritz T. 2010. Changes in diurnal patterns within the *Populus* transcriptome and metabolome in response to photoperiod variation. Plant, Cell and Environment 33: 1298–1313.

Ivanov A.S., Zgoda V.G., Archakov A.I. 2011. Technologies of protein interactomics: a review. Russian Journal of Bioorganic Chemistry 37: 4–16.

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

Janz D., Behnke K., Schnitzler J.P., Kanawati B., Schmitt-Kopplin P., Polle A. 2010. Pathway analysis of the transcriptome and metabolome of salt sensitive and tolerant poplar species reveals evolutionary adaption of stress tolerance mechanisms. BMC Plant Biology 10: Article Number 150.

Jorrín-Novo J.V., Maldonado A.M., Echevarría-Zomeño S., Valledor L., Castillejo M.A., Curto M., Valero J., Sghaier B., Donoso G., Redondo I. 2009. Plant proteomics update (2007–2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. Journal of Proteomics 72: 285–314.

Kerner R., Winkler J.B., Dupuy J.W., Jürgensen M., Lindermayr C., Ernst D., Müller-Starck G. 2011. Changes in the proteome of juvenile European beech following three years exposure to free-air elevated ozone. iForest Biogeoscience and Forestry 4: 69–76.

Kersten B., Agrawal G.K., Iwahashi H., Rakwal R. 2006. Plant phosphoproteomics: a long road ahead. Proteomics 6: 5517–5528.

Kim H.K., Choi Y.H., Verpoorte R. 2011. NMR-based plant metabolomics: where do we stand, where do we go? Trends in Biotechnology 29: 267–275.

Kitano H. 2002. Systems biology: a brief overview. Science 295: 1662–1664.

Kohler A., Delaruelle C., Martin D., Encelot N., Martin F. 2003. The poplar root transcriptome: analysis of 7000 expressed sequence tags. FEBS Letters 542: 37–41.

Lo S.H. 2007. Reverse Interactomics: From Peptides to Proteins and to Functions. ACS Cell Biology 2: 93–95.

McEntyre J., Ostell J. 2002. The NCBI Handbook. Bethesda (MD): National Center for Biotechnology Information (US).

Medina I, Carbonell J., Pulido L., Madeira S.C., Goetz S., Conesa A., Tárraga J., Pascual-Montano A., Nogales-Cadenas R., Santoyo J., García F., Marbá M., Montaner D., Dopazo J. 2010. Babelomics: an integrative platform for the analysis of transcrip-

tomics, proteomics and genomic data with advanced functional profiling. Nucleic Acids Research 38: 210–213.

Miernyk J.A., Hajduch M. 2011. Seed proteomics. Journal of Proteomic 74: 389 – 400.

National Center for Biotechnology Information Entrez Genome Project, http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html.

National Center for Biotechnology Information, Expressed Sequence Tag databases, http://www.ncbi.nlm.nih.gov/dbEST/dbEST.

Oeljeklaus S., Meyer H.E, Warscheid B. 2009. Advancements in plant proteomics using quantitative mass spektrometry. Journal of Proteomics 72: 545–554.

Ohlrogge J., Benning C. 2000. Unraveling plant metabolism by EST analysis. Current Opinion in Plant Biology 3: 224–228.

Palumbo A.M., Smith S.A., Kalcic C.L., Dantus M., Stemmer P.M., Reid G.E. 2011. Tandem mass spectrometry strategies for phosphoproteome analysis. Mass Spectrometry Reviews 30: 600–625.

Pawłowski T., Kalinowski A. 2003. Qualitative and quantitative changes in proteins in *Acer platanoides* L. seeds during maturation. Acta Biologica Cracoviensia Series Botanica 72: 289–293.

Pawłowski T.A. 2007. Proteomics of European beech (*Fagus sylvatica* L.) seed dormancy breaking: Influence of abscisic and gibberellic acids. Proteomics 7: 2246–2257.

Pawłowski T.A. 2009. Proteome analysis of Norway maple (*Acer platanoides* L.) seeds dormancy breaking and germination: influence of abscisic and gibberellic acids. BMC Plant Biology 9: Article Number 48.

Pawłowski T.A. 2010. Proteomic approach to analyze dormancy breaking of tree seeds. Plant Molecular Biology 73: 15–25.

Peck S.C. 2006. Phosphoproteomics in Arabidopsis: moving from empirical to predictive science. Journal of Experimental Botany 57: 1523–1527.

Peńa L., Séguin A. 2001. Recent advances in the genetic transformation of trees. Trends in Biotechnology 19: 500–506.

Plomion C., Lalanne C., Claverol S., Meddour H., Kohler A., Bogeat-Triboulot M.B., Barre A, Le Provost G., Dumazet H., Jacob D., Bastien C., Dreyer E., de Daruvar A., Guehl J.M., Schmitter J.M., Martin F., Bonneu M. 2006. Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. Proteomics 6: 6509–6527.

Remmerie N., Vijlder T., Laukens K., Dang T.H., Lemière F., Mertens I., Valkenborg D., Blust R., Witters E. 2011. Next generation functional proteomics in non-model plants: a survey on techniques and applications for the analysis of protein complexes and post-translational modifications. Phytochemistry 72: 1192–1218.

Rigault P., Boyle B., Lepage P., Cooke J.E.K., Bousquet J., MacKay J.J. 2011. A white spruce gene catalog for conifer genome. Plant Physiology 157: 14–28.

Robinson A.R., Gheneim R., Kozak R.A., Ellis D.D., Mansfield S.D. 2005. The potential of metabolite profiling as a selection tool for genotype discrimination in *Populus*. Journal of Experimental Botany 56: 2807–2819.

Robinson A.R., Mansfield S.D. 2009. Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. The Plant Journal 58: 706–714.

Qiu Q., Ma T., Hu Q., Liu B., Wu Y., Zhou H., Wang Q., Wang J., Liu J. Sederoff R. 2011. Genome-scale transcriptome analysis of the desert poplar, *Populus euphratica*. Tree Physiology 31: 452–461.

Schenk M.F., Cordewenera J.H.G., America A.H.P., Peters J., Smulders M.J.M., Gilissen L.J.W.J. 2011. Proteomic analysis of the major birch allergen Bet v 1 predicts allergenicity for 15 birch species. Journal of Proteomics 74: 1290–1300.

Sergeant K., Spiess N., Renaut J., Wilhelm E., Hausman J. F. 2011. One dry summer: a leaf proteome study on the response of oak to drought exposure. Journal of Proteomics 74: 1385–1395.

Sjödin A., Bylesjö M., Skogström O., Eriksson D., Nilsson P., Rydén P., Jansson S. Karlsson J. 2006. UPSC-BASE – *Populus* transcriptomics online. Plant Journal 48: 806–817.

Sjödin A. 2007. Populus transcriptomics – from noise to biology. PhD thesis. Plant Science Centre Department of Plant Physiology Umeå University, Umeå, Sweden.

Song D., Shen J., Li L. 2010. Characterization of cellulose synthase complexes in *Populus* xylem differentiation. New Phytologist 187: 777–790.

Stapley J., Reger J., Feulner P.G., Smadja C., Galindo J., Ekblom R., Bennison C., Ball A.D., Beckerman A.P., Slate J. 2010. Adaptation genomics: the next generation. Trends in Ecology and Evolution 25: 705–712.

Sterky F., Regan S., Karlsson J., Hertzberg M., Rohde A., Holmberg A., Amini B., Bhalerao R., Larsson M., Villarroel R., Montagu M.V., Sandberg G.R., Olssoni O., Teeri T.T., Boerjan W., Gustafsson P., Uhlén. M., Sundberg B., Lundeberg J. 1998. Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags. Proceedings of the National Academy of Sciences of the United States of America 95: 13330–13335.

Sterky F., Bhalerao R.R., Unneberg P., Segerman B., Nilsson P., Brunner A.M., Charbonnel-Campaa L., Lindvall J.J., Tandre K., Strauss S.H., Sundberg B., Gustafsson P., Uhlén M., Bhalerao R.P., Nilsson O., Sandberg G., Karlsson J., Lundeberg J., Jansson S. 2004. A *Populus* EST resource for plant functional genomics. Proceedings of the National Academy of Sciences of the United States of America 101: 13951–13956.

Street N.R., James T.M., James T., Mikael B., Jaakko K., Mark B., Taylor G. 2011. The physiological, transcriptional and genetic responses of an ozone- sensitive and an ozone tolerant poplar and selected extremes of their F2 progeny. Environmental Pollution 159: 45–54.

Surget-Groba Y., Montoya-Burgos J.I. 2010. Optimization of de novo transcriptome assembly from next-generation sequencing data. Genome Research 20: 1432–1440.

Szczotka Z., Pawłowski T., Krawiarz K. 2003. Proteins and polyamines relation during dormancy breaking of European beech (*Fagus sylvatica* L.) seeds. Acta Physiologiae Plantarum 25: 423–435.

Taylor G. 2002. *Populus*: Arabidopsis for forestry. Do we need a model tree? Annals of Botany 90: 681–689.

Tonge R., Shaw J., Middleton B., Rowlinson R., Rayner S., Young J., Pognan F., Hawkins E., Currie I., Davison M. 2001. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 1: 377–396.

Tuskan G.A., DiFazio S.P., Teichmann T. 2003. Poplar genomics is getting popular: the impact of the poplar genome project on tree research. Plant Biology 6: 2–4.

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., Déjardin 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., Leplé 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., Rouzé 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 Peer Y., Rokhsar D. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313: 1596–1604.

Ueno S., Provost G., Léger V., Klopp C., Noirot C., Frigerio J.M., Salin F., Salse J., Abrouk M., Murat F., Brendel O., Derory J., Abadie P., Léger P., Cabane C., Barré A., de Daruvar A., Couloux A., Wincker P., Reviron M.P., Kremer A., Plomion C. 2010. Bioinformatic analysis of ESTs collected by Sanger and pyrosequencing methods for a keystone forest tree species: oak. BMC Genomics 11: Article Number 650.

Valero Galván J., Valledor L., Navarro Cerrillo R.M., Gil Pelegrín E., Jorrín-Novo J.V. 2011. Studies of variability in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.) through acorn protein profile analysis. Journal of Proteomics 74: 1244–1255.

Vera J.C., Wheat C.W., Fescemyer H.W., Frilander M.J., Crawford D.L., Hanski I., Marden J.H. 2008. Rapid transcriptome characterization for a non model organism using 454 pyrosequencing. Molecular Ecology 17: 1636–1647.

Ward J.L., Baker J.M., Beale M.H. 2007. Recent applications of NMR spectroscopy in plant metabolomics. The FEBS Journal 274: 1126–1131.

Ward J.L., Baker J.M., Miller S.J., Deborde C., Maucourt M., Biais B., Rolin D., Moing A., Moco S., Vervoort J., Lommen A., Schäfer H., Humpfer E., Beale M.H. 2010. An inter-laboratory comparison demonstrates that [^{1}H]-NMR metabolite fingerprinting is a robust technique for collaborative plant metabolomic data collection. Metabolomics 6: 263–273.

Wienkoop S., Morgenthal K., Wolschin F., Scholz M., Selbig J., Weckwerth W. 2008. Integration of metabolomic and proteomic phenotypes. Analysis of data covariance dissects starch and RFO metabolism from low and high temperature compensation response in *Arabidopsis thaliana*. Molecular and Cellular Proteomics 7: 1725–1736.

Wishart D.S. 2011. Advances in metabolite identification. Bioanalysis 3: 1769-1782.

Zhang J., Gao G., Chen J.J., Taylor G., Cui K.M., He X.Q. 2011. Molecular features of secondary vascular tissue regeneration after bark girdling in *Populus*. New Phytologist 192: 869–884.