

POSTER PRESENTATION

Open Access

# The transcriptome of a *Populus* pseudo-backcross identifies genes and pathways co-expressed with monolignol biosynthesis

Evandro Novaes<sup>1\*</sup>, Carolina Boaventura-Novaes<sup>2</sup>, Alexandre Coelho<sup>1</sup>, Matias Kirst<sup>3</sup>

From IUFRO Tree Biotechnology Conference 2011: From Genomes to Integration and Delivery  
Arraial d'Ajuda, Bahia, Brazil. 26 June - 2 July 2011

## Background

Lignin is the main impediment for efficient cellulose extraction and ethanol production from woody tissues. Even though, the biosynthesis of the lignin constituents - the monolignols - is well characterized, little is known about the genetic control of the natural variation in lignin content and composition. The few association studies performed to date in forest species only uncover 5-20% of the heritable variation of quantitative traits such as wood composition. The “missing heritability” can be explained in part by the low resolution of these pioneering association studies in forest species. However, as observed in genome-wide association studies with humans, a large proportion of the “missing heritability” is likely to occur due to other factors, such as the abundance of rare alleles observed in forest species and complex epistatic interaction between genetic elements. Causal rare alleles and genetic interactions remain undetectable with current statistical methods. In order to shed light on possible interactions between lignin biosynthesis genes and other pathways, we analyzed the transcriptome of 181 genotypes of a pseudo-backcross family of *Populus*. The analyses allowed identification of genes and pathways that were highly co-expressed with genes involved in the biosynthesis of monolignols. Correlations at the transcript level offer initial evidence of possible interactions between genetic elements for the production of monolignols.

## Methods

A previously published microarray data [1] was utilized. Briefly, this microarray contains a gene expression probe for every gene annotated in the genome of *Populus trichocarpa* (v1.1.). Microarray chips were hybridized with cDNA synthesized from xylem tissue of 181 genotypes from a pseudo-backcross pedigree of *Populus trichocarpa* x *P. deltoides*. This family was also analyzed for lignin content on xylem tissue [2]. Genetic differences in gene expression allowed us to correlate the transcript abundance of 23 previously identified xylem-specific monolignol biosynthetic-genes [3] against the cDNA levels of all annotated genes of *Populus*. Correlation was estimated based on Spearman's rank method in R. The top correlated genes ( $r > 0.75$ ,  $p$ -value  $< 0.001$ ) were clustered with the 23 monolignol biosynthetic-genes based on a “Modulated Modularity Clustering” method [4]. The genes clustered with the 23 monolignol biosynthetic-genes were GO annotated based on the BlastP best hit against the *Arabidopsis* gene models. A Fisher's exact test was applied to identify GO terms enriched within these clusters.

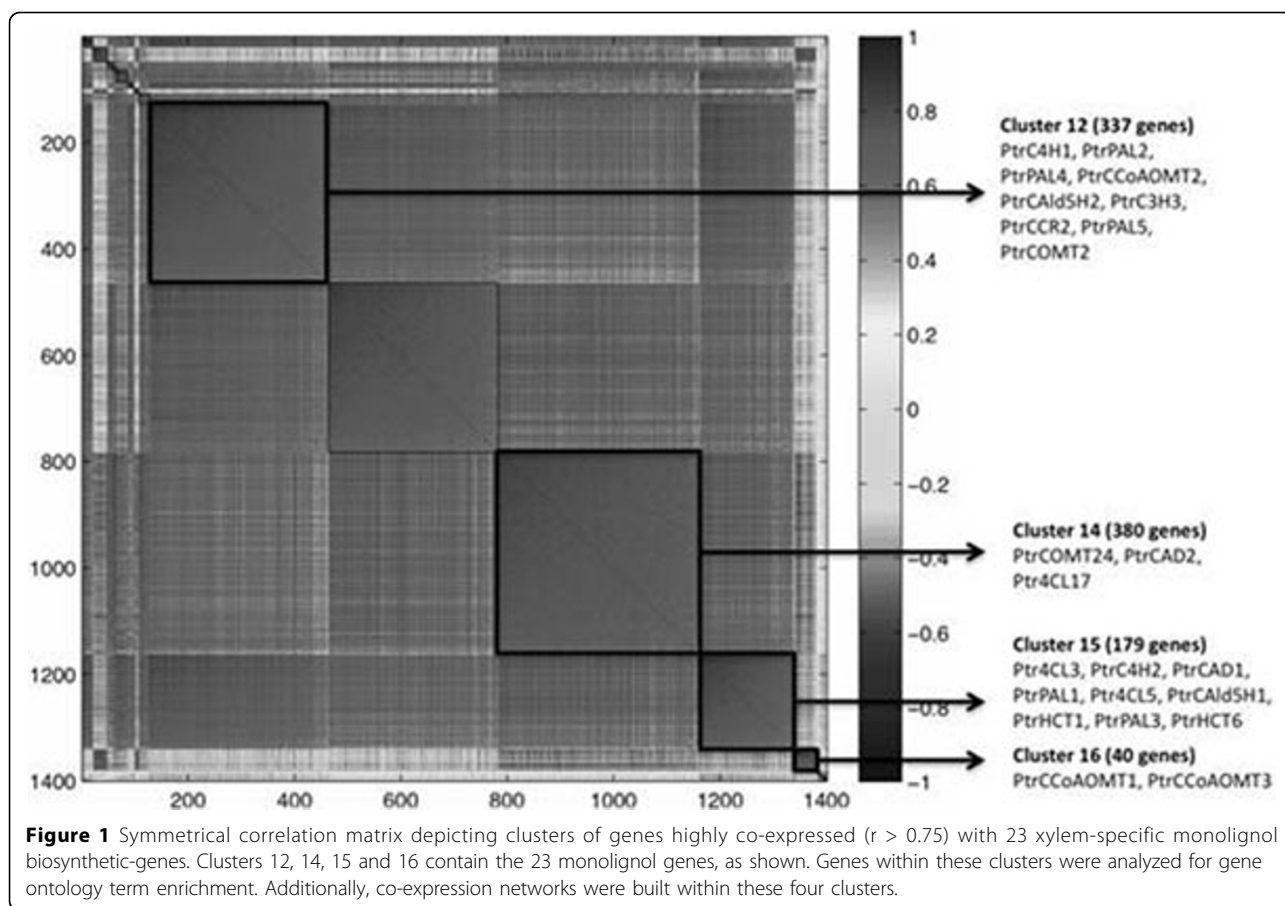
## Results

Utilizing a correlation threshold of 0.75 ( $p$ -value  $< 0.001$ ), 1369 genes were correlated with at least one of the 23 monolignol biosynthetic-genes. These genes were clustered together with the lignin phenotypic trait and 78 genes that were correlated with lignin ( $r > 0.45$ ). The 23 monolignol genes were clustered in four groups, containing a total of 936 genes (Figure 1). Surprisingly, the lignin trait itself did not cluster with any other gene, and the vast majority (75%) of the top 78 lignin-correlated genes were grouped on clusters 12 and 15. Clusters 14 and 16 did not have lignin-correlated genes. We

\* Correspondence: [novaes@agro.ufg.br](mailto:novaes@agro.ufg.br)

<sup>1</sup>Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Caixa Postal 131, CEP 74690-900, Goiânia, Brazil

Full list of author information is available at the end of the article



analyzed the GO annotation among genes within the four monolignol-related clusters. A Fisher's exact test was utilized to find GO terms enriched within these clusters when compared to all annotated genes in the genome of *Populus*. Sixty-nine GO terms were significantly enriched ( $FDR < 0.01$ ) among genes clustered with the 23 monolignol genes: 11 terms are from the ontology Cellular Component (CC), 26 from Molecular Function (MF) and 32 from Biological Process (BP). As expected, most of these MF and BP terms (>65%) are related to lignin, cellulose and hemicellulose biogenesis. This high specificity indicates that even the genes with unknown MF and BP (260 of the 936 monolignol clustered genes) might be involved in cell-wall biogenesis. Enriched GO terms not directly related to cell-wall biogenesis include "drought-recovery", "salicylic acid catabolism", "response to cadmium ion" and "response to zinc ion". In addition to the significantly enriched GO terms, we also constructed a co-expression network including the core, most highly correlated genes ( $r > 0.85$ ) within each of the four identified, monolignol biosynthetic clusters. These networks, representing the most likely interacting genes, will be presented.

## Conclusion

This work presents a set of approximately 900 genes highly co-expressed with xylem-specific monolignol biosynthesis genes. Many of these genes are currently annotated with unknown function, or are not known to be involved in cell-wall biogenesis. We offer initial evidence towards a role or interaction of these genes in the biosynthesis of lignin and possibly other cell-wall components.

## Author details

<sup>1</sup>Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Caixa Postal 131, CEP 74690-900, Goiânia, Brazil. <sup>2</sup>Programa de Pós-Graduação em Genética e Melhoramento de Plantas, Universidade Federal de Goiás, Caixa Postal 131, CEP 74690-900, Goiânia, Brazil. <sup>3</sup>School of Forest Resources and Conservation, University of Florida, PO Box 110410, Gainesville, USA.

Published: 13 September 2011

## References

1. Drost DR, Benedict CI, Berg A, Novaes E, Novaes CRDB, Yu Q, Miles BL, Dervinis C, Maia JM, Yap J, Miles B, Kirst M: **Diversification in the genetic architecture of gene expression and transcriptional networks in organ differentiation of *Populus***. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(18):8492-8497.

2. Novaes E, Osorio L, Drost DR, Miles BL, Boaventura-Novaes CRD, Benedict C, Dervinis C, Yu Q, Sykes R, Davis M, Martin TA, Peter GF, Kirst M: **Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels.** *New Phytologist* 2009, **182**(4):878-890.
3. Shi R, Sun YH, Li QZ, Heber S, Sederoff R, Chiang VL: **Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes.** *Plant and Cell Physiology* 2010, **51**(1):144-163.
4. Stone EA, Ayroles JF: **Modulated Modularity Clustering as an Exploratory Tool for Functional Genomic Inference.** *PLoS Genet* 2009, **5**(5):e1000479.

doi:10.1186/1753-6561-5-S7-P119

**Cite this article as:** Novaes *et al.*: The transcriptome of a *Populus* pseudo-backcross identifies genes and pathways co-expressed with monolignol biosynthesis. *BMC Proceedings* 2011 **5**(Suppl 7):P119.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

