Massive transcriptome changes during leaf senescence in field grown *Populus trichocarpa* Nisqually-1 using KBase tools

Haiwei Lu, Vindhya Amarasinghe, and Steven H. Strauss
Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA
Steve.Strauss@oregonstate.edu

Abstract

Leaf senescence is the programmed death of foliage that is essential to plant reproduction and survival in temperate environments. In this study, we examined transcriptome changes during leaf senescence under natural conditions in *Populus trichocarpa* using RNA sequencing (RNA-seq). Using RNA from leaf samples collected at the end of each month from May to October in 2012, 2015, and 2016, we built and sequenced a total of 54 RNA sequencing (RNA-seq) libraries, each representing one time point and containing three biological replicates. We applied the HISAT2-StringTie-DESeq2 pipeline (enabled by Kbase) to map and quantify the short reads and identify differentially expressed genes (DEGs). We found that each library had at least 18 million reads mapped to the reference genome. Pair-wise comparisons revealed extensive changes in gene expression. For example, of 32,328 *Populus* gene models analyzed using DESeq2, we found that 14,962 (46.3%) were differentially expressed between May 2015 and October 2015, and 9,202 (28.5%) were differentially expressed between July 2015 and September 2015 (FDR < 0.05). Plotting expression profiles of selected genes revealed reduced chloroplast activity and chlorophyll biosynthesis in August. A number of gene ontology (GO) terms, such as those related to metabolic process, reproduction, signaling, and cellular process, were enriched in early senescence (*i.e.*, July-August comparison). Additional analysis, such as hierarchical clustering, identification of novel isoforms and *cis*-acting elements, and co-expression network analysis, is underway.

Goals

- Examine gene expression patterns in depth during temperature and photoperiod-induced natural senescence in *Populus* in depth.
- Develop a toolkit useful for metabolic engineering of senescence related traits, such as knowledge of key transcription factors, regulatory networks, and promoters useful for timed accumulation of biopolymers.

Hypotheses to Test

- Samples collected from different months differ in expression patterns, and cluster according to chronology and physiological state
- Genes related to chlorophyll biosynthesis, photosynthesis, protein synthesis and other energy-requiring activities, are down-regulated as leaves senesce
- Genes related to reactive oxygen species (ROS) and catalytic activity are up-regulated as leaves senesce
- Due to the depth of sampling in our RNAseq study compared to prior work, we will detect large numbers of genes not previously associated with senescence in perennial plants

Experimental Strategy

Collected leaf samples from *Populus trichocarpa* at the end of each month from May to October in 2012, 2015, and 2016

Monitored changes in senescence-related physiological parameters

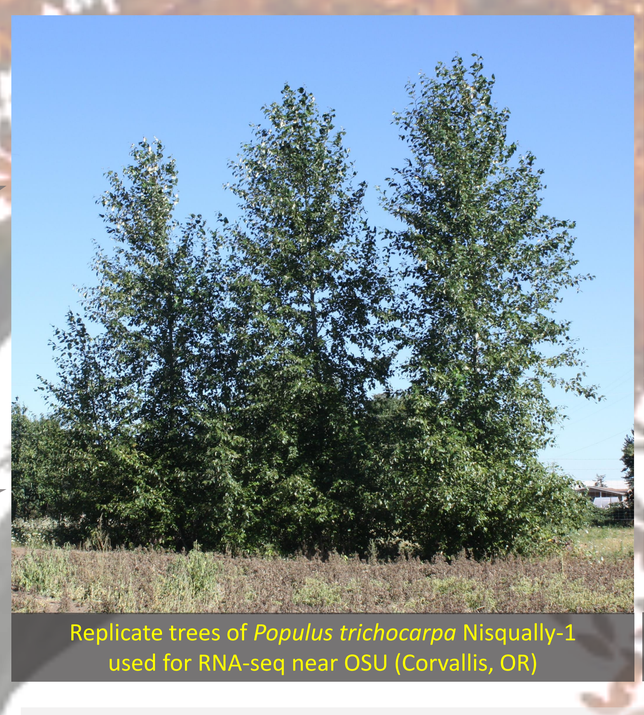
Built and sequenced a total of 54 RNA-seq libraries (three biological replicates for each collection timepoint)

Mapped RNA-seq reads to the reference genome using HISAT2

Assembled and quantified transcripts using StringTie

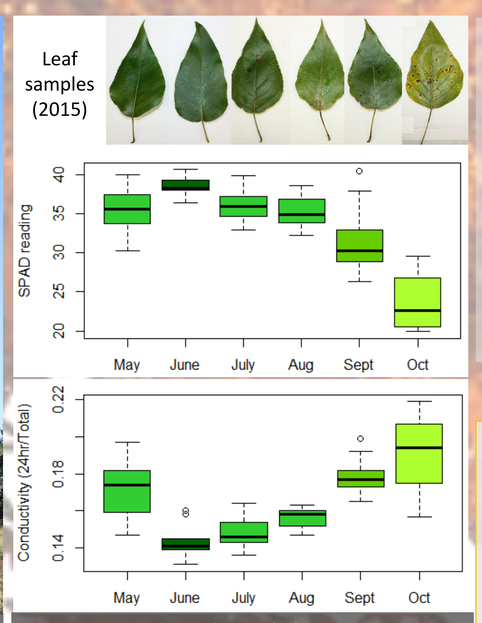
Identified DEGs (FDR < 0.05) using DESeq2

Performed GO enrichment using AgriGO v2



Replicate trees of *Populus trichocarpa* Nisqually-1 used for RNA-seq near OSU (Corvallis, OR)

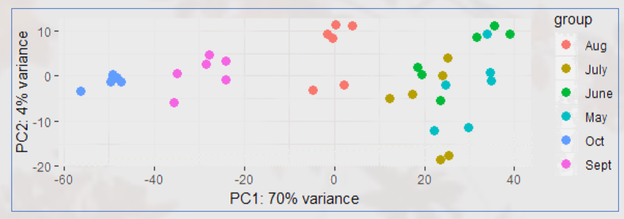
Results



Extensive physiological changes among collections (SPAD is an index of chlorophyll)

Changes in three senescence-related physiological parameters, chlorophyll content (indicated by SPAD reading), photochemical efficiency of PSII, and membrane ion leakage (indicated by conductivity), were monitored at each collection timepoint in 2015 and 2016. For each of the parameters, we compared the difference among collections using ANOVA and Tukey HSD. Based on 2015 data, chlorophyll content reached highest in June, and decreased significantly ($p < 0.05$) in Sept and Oct. Meanwhile, Sept and Oct samples had significantly higher conductivity ($p < 0.05$) compared with samples from the other months. Taken together, our collections were not merely senescence focused, rather they covered several different developmental states, which likely include active growth, maturation, early senescence, and late senescence.

Collections clustered into four chronological groups

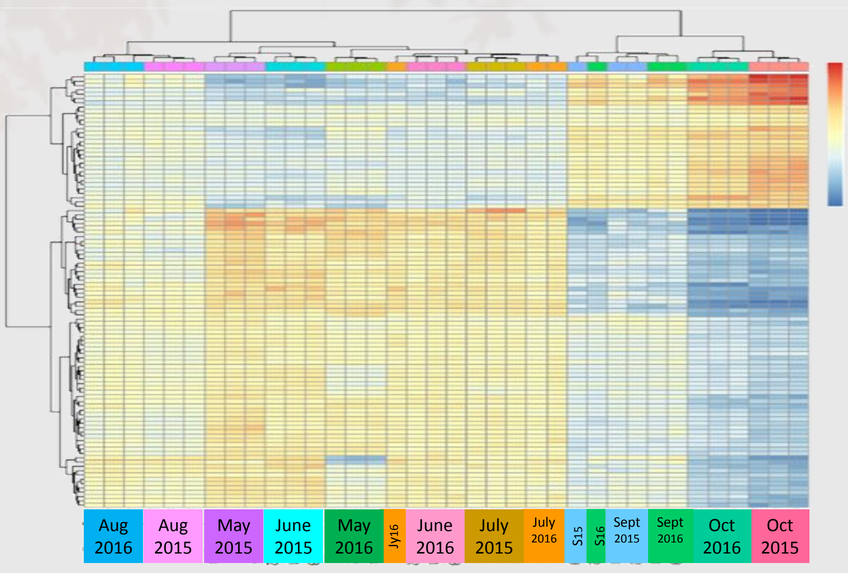


Principal Components Analysis showed that the time of the growing season was the main determinant of expression patterns, and was consistent among two years of study

Because 2012 sequencing data was produced a while ago using a different platform, we focused our analysis on 2015 and 2016 data. On average, each library had 24 million mapped reads mapped to the *P. trichocarpa* genome (v3.1; Phytozome), with 21 million reads mapped once and 3 million reads mapped more than once. A total of 42,950 transcripts were assembled and quantified by StringTie. After removing genes with low counts, 30,297 genes were used as DESeq2 input for DEG identification. Results from 2015 data are shown below.

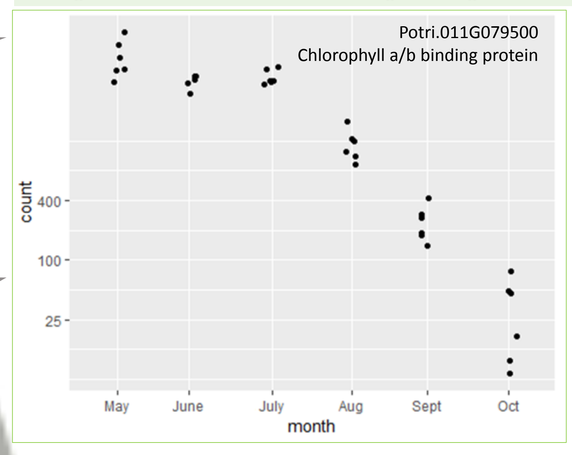
Month - Year	June2015	July2015	Aug2015	Sept2015	Oct2015
May2015	1,739 (5.4%) • 988 (3.1%) • 751 (2.3%)	3,033 (9.4%) • 1,521 (4.7%) • 1,512 (4.7%)	5,679 (17.6%) • 2,945 (9.1%) • 2,734 (8.5%)	10,636 (32.9%) • 5,406 (16.7%) • 5,230 (16.2%)	14,962 (46.3%) • 7,549 (23.4%) • 7,413 (22.9%)
June2015		2,276 (7.0%) • 901 (2.8%) • 1,375 (4.3%)	5,956 (18.4%) • 2,915 (9.0%) • 3,041 (9.4%)	12,049 (37.3%) • 6,203 (19.2%) • 5,846 (18.1%)	15,434 (47.7%) • 7,842 (24.3%) • 7,592 (23.5%)
July2015			3,875 (12.0%) • 1,985 (6.1%) • 1,890 (5.8%)	9,202 (28.5%) • 4,738 (14.7%) • 4,464 (13.8%)	13,877 (42.9%) • 7,153 (22.1%) • 6,724 (20.8%)
Aug2015				4,128 (12.8%) • 2,027 (6.3%) • 2,101 (6.5%)	10,897 (33.7%) • 5,632 (17.4%) • 5,265 (16.3%)
Sept2015					7,226 (22.4%) • 3,922 (12.1%) • 3,304 (10.2%)

Total No. (%)
- Down regulated No. (%)
- Up regulated No. (%)

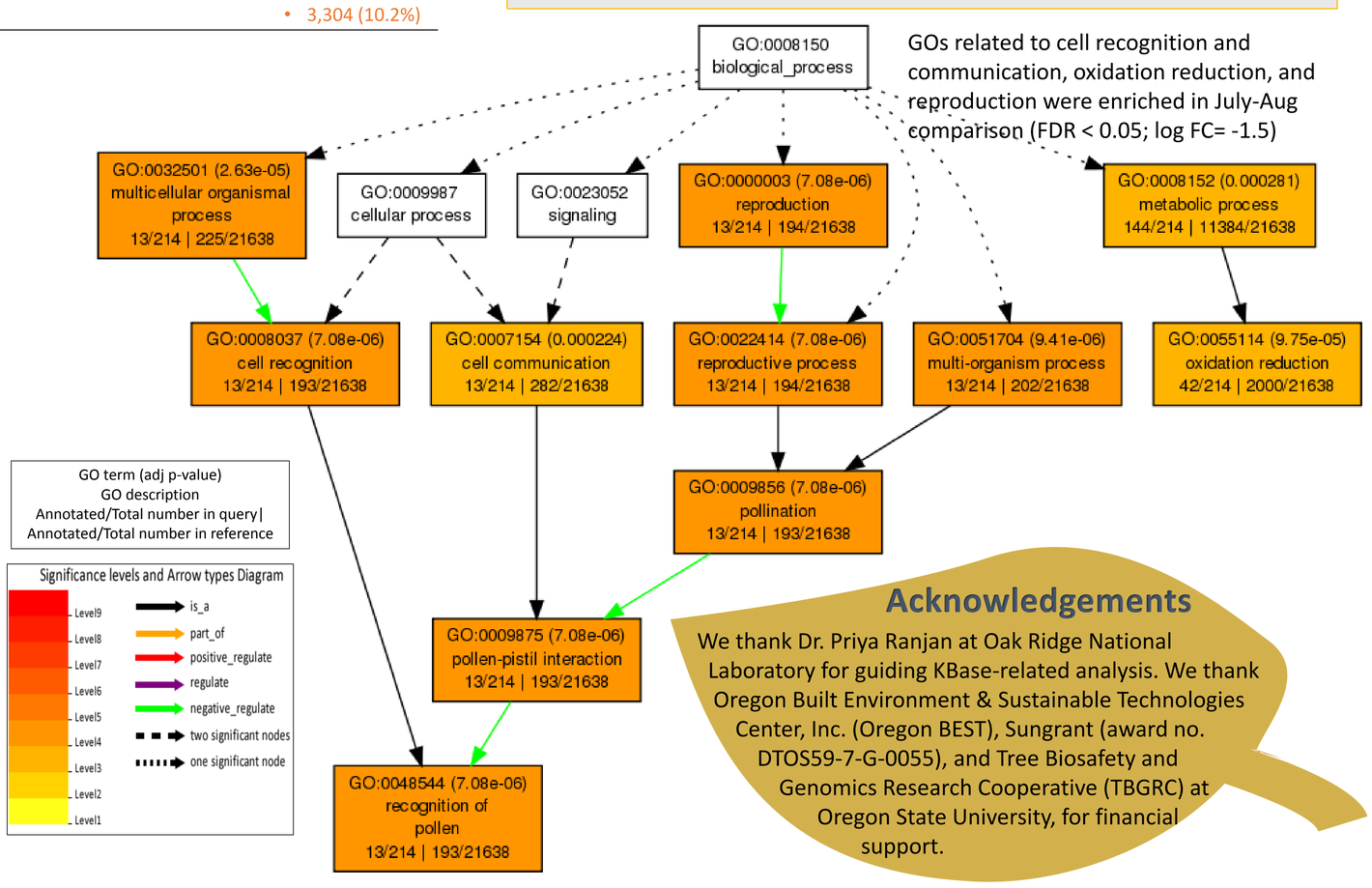


Heatmap (based on top 100 genes in the May 2015 vs Oct 2015 comparison) clustered samples into four groups, agreeing with the result from Principal Components Analysis.

Expression of chloroplast/chlorophyll related genes was reduced starting in August



GOs related to cell recognition and communication, oxidation reduction, and reproduction were enriched in July-Aug comparison (FDR < 0.05; log FC= -1.5)



Future Directions

- Perform clustering analysis to group collections into different developmental states (*e.g.*, growth, maturation, early senescence, and late senescence)
- Refine DEGs and GOs (after assigning collections into development states)
- Group DEGs into contrasting gene expression groups
- Identify novel isoforms associated with senescence (enabled by StringTie, the assembler used in the analysis pipeline)
- Identify sequence motifs associated with each gene expression group
- Validate gene expression using 2012 data and qRT-PCR

Acknowledgements

We thank Dr. Priya Ranjan at Oak Ridge National Laboratory for guiding KBase-related analysis. We thank Oregon Built Environment & Sustainable Technologies Center, Inc. (Oregon BEST), Sungrant (award no. DTOS59-7-G-0055), and Tree Biosafety and Genomics Research Cooperative (TBGRC) at Oregon State University, for financial support.