The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Transcriptome dynamics in the dormancy-spring growth transition of Douglas-fir needles

Brian J. Knaus[1]   Peter C. Dolan[2]   Dee Denver[3]   Rich Cronn[1]

[1]Pacific Northwest Research Station, USDA Forest Service

[2]Department of Mathematics, University of Minnesota, Morris

[3]Department of Zoology, Oregon State University

Forest Tree Workshop, PAG XX
Sunday, January 15th, 2012

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

1. The Douglas-fir needle transcriptome.

2. Differential expression: the dormancy-spring growth transition.

3. Annotation: are differentially expressed isogroups annotated?

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

- Huge geographic and environmental range
- 22 million hectares in the US/Canada
- Planted in Europe, New Zealand, Chile
- 8 billion board feet of lumber in 2002
- Large breeding programs - more than: 4 million progeny from 34,000 parents on 1,000 progeny test sites

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
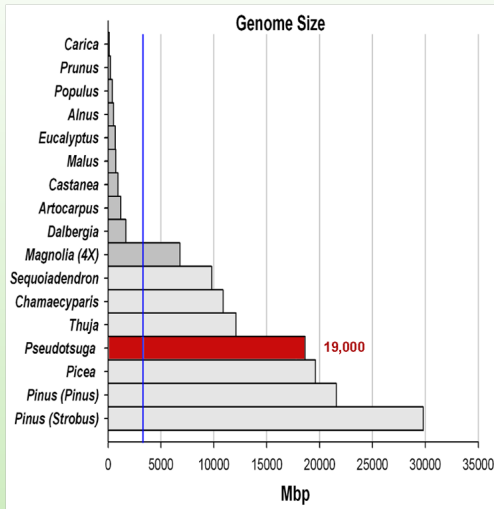Annotation: are differentially expressed isogroups annotated?

Adaptive genetic variation in Douglas-fir
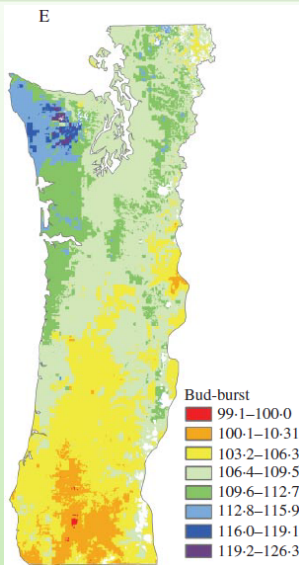
- Strong correlation between climate and quantitative traits.
    - Spring bud burst
    - Fall cold hardiness
    - Growth rhythm, biomass accumulation
    - Trunk taper
- Association is basis for seed transfer guidelines.
- Association is basis for candidate gene tests for adaptive variation.

St Clair, J.B., Mandel, N.L. and Vance-Borland, K.W. 2005.
Genecology of Douglas fir in western Oregon and Washington.
Annals of Botany 96(7):1199-1214.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Challenges to conifer genomics projects

- Genome size much larger than current model organisms
- Douglas-fir is average-sized for a conifer.
- Conifers are incompletely characterized at a genic level.
- Uncertainty in evolutionary, functional homology to 'candidate' genes.
- Adaptations unique to Douglas-fir.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

RNA-Seq methodology

Wet chemistry

- Urea/LiCl total RNA.
- Illumina library preparation.
  - Poly(A)+ selection
  - Tru-Seq kit
  - Strand-specific modification, dUTP method (Parkhomchuk, 2009)
- Illumina GAII sequencing.
  - 100 bp single-end sequences
  - 6-plex index pools

Bioinformatics

- *De novo* assembly with Trinity.
  - RNA-Seq assembler
  - Strand specific
  - Isogroup aware
- Validation.
  - blastn
  - tblastx
  - Custom perl and R

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

The Douglas-fir transcriptome

De novo transcriptome assembly

Increase	v2012[1]	v2011[2]	
166X	500	3	Million reads
2.1X	112.8	53.6	Assembled million bp
3.6X	90,086	25,002	Isogroups
-	-	102,623	Singletons
3.7X	143,392	38,589[3]	Isoforms

[1]Libraries derived from needle tissue, sequenced with Illumina technology, assembled with Trinity.
[2]Libraries derived from multiple tissues, sequenced with Roche/454 and Sanger technologies, assembled with Newbler.
[3]Excludes singletons.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Core eukaryotic gene (KOG) validation

KOGs tblastn against Douglas-fir needle transcriptome

112	Zero hits
105	**One hit**
101	Two hits
50	Three hits
90	More than three hits
458	Total KOG genes

75.5% of *A. thaliana* KOG genes appear to occur one or more times as (nearly) fully assembled.

Where a hit is an HSP with an *e*-value less than e^{-10} and the alignment is at least 90% as long as the KOG sequence (a subglobal alignment).

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

The Douglas-fir needle meta-genome
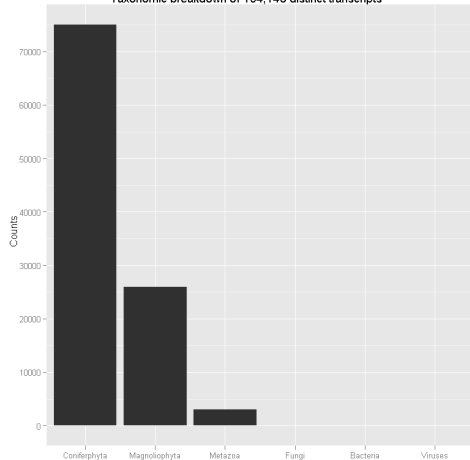
72.6% of transcripts were successfully assigned a taxonomic annotation.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Experimental design
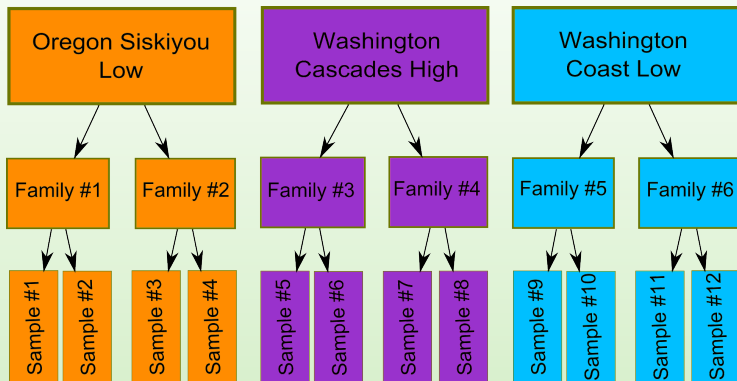
Two half-sib families nested within three seed sources for a total of 12 samples. (Three 32 sample families collected during 2011 - final sequencing dependent on sequencing cost.)

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?
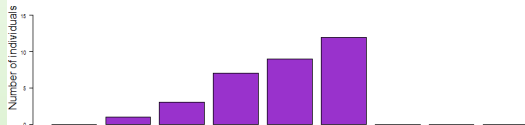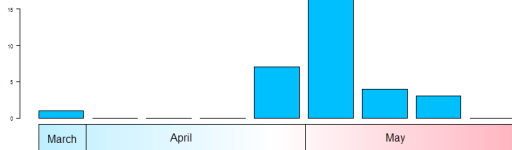
Phenotypic differences through time:

March	Dormant
April	Budburst
May	Physiologically active

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Differential expression methodology

Read mapping

- Same reads used from transcriptome assembly.
  - RNA-Seq
  - Illumina GAII
- Reads mapped with bowtie.
- Summarize into count tables with custom perl.
- Summarize isoforms to isogroups.
  - Mitigates multimap reads

Hypothesis testing

- Negative binomial GLMs.
- Library size normalized by random subset.
- Negative binomial dispersion parameters estimated in EdgeR (tag-wise, trend based).
- Control (March) contrasts.
- Correction for multiple comparisons (*q*-values) followed the method of Benjamini and Hochberg (1995).
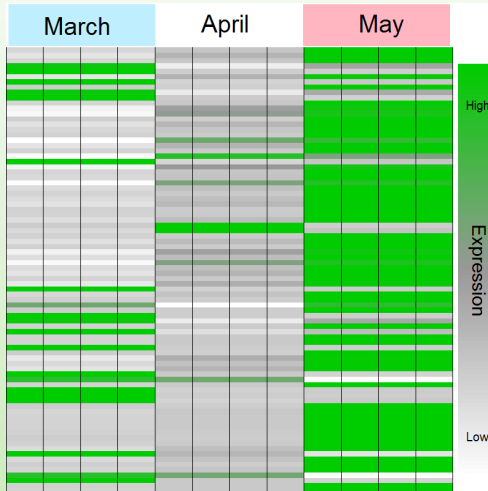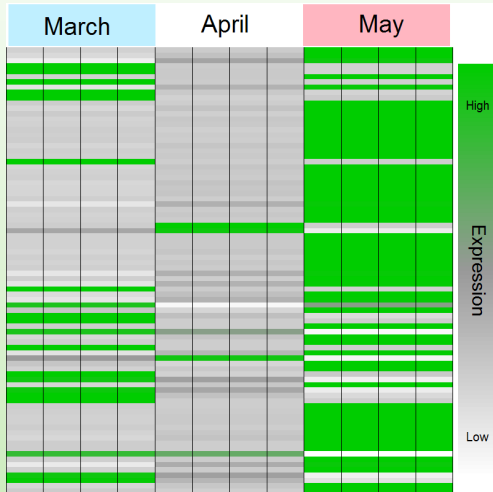
The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Differentially expressed isogroups
Difference in mean expression during the dormancy-spring growth transition
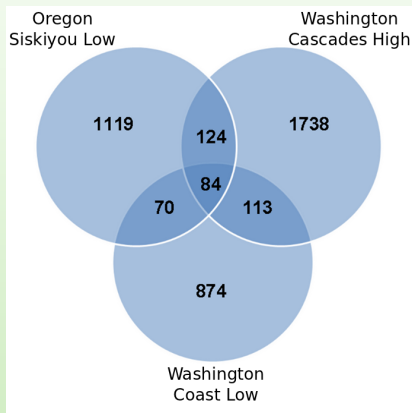
- Hundreds to thousands of differentially expressed isogroups.
- 84 isogroups common to all three seed sources.
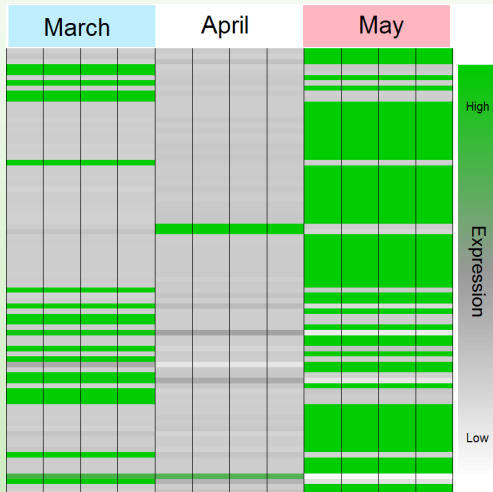- 391 isogroups differentially expressed by two seed sources.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Oregon Siskiyous, Low Elevation
84 common D.E. isogroups

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

# Washington Cascades, High Elevation
84 common D.E. isogroups

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Washington Coast, Low Elevation
84 common D.E. isogroups

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Washington Coast, Low Elevation
84 common D.E. isogroups, raw data

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
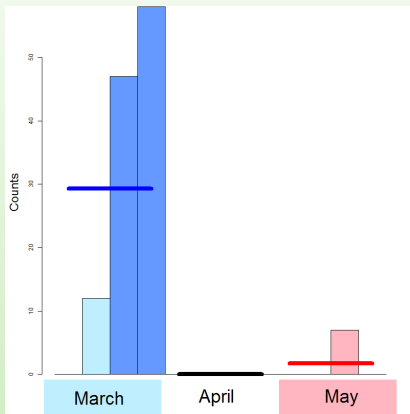Annotation: are differentially expressed isogroups annotated?
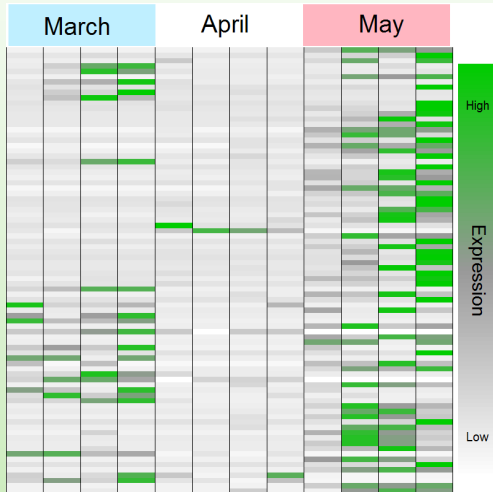
# Washington Coast, Low Elevation
Isogroup 81, raw data


- Fitted values do not account for family differences.
- Larger sample size may allow partitioning of variance at the level of family.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Taxonomic annotation
84 common D.E. isogroups

blastn

blastx

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
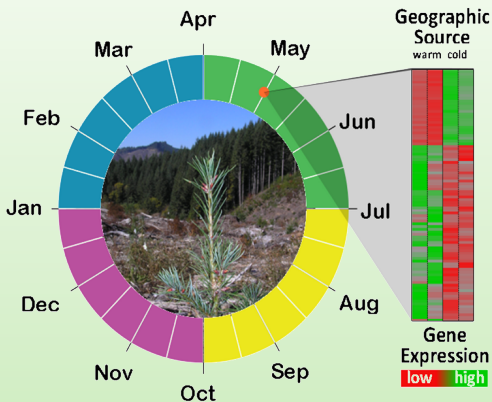Annotation: are differentially expressed isogroups annotated?
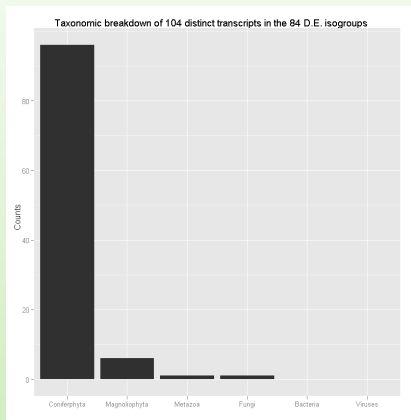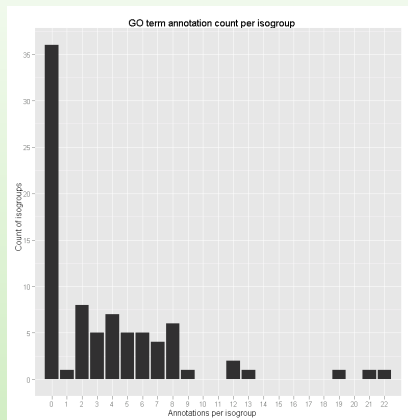
Summary

- *De novo* assembly of transcriptomes using short read technologies ($\approx$ 100 bp) is possible.
- Some isogroups appear more abundant in the physiologically dormant state (dormancy genes).
- Annotation indicates that 57.1% these isogroups may be homologous to annotated genes. The remaining 43.0% of isogroups may be novel to Douglas-fir (or conifers in general).

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

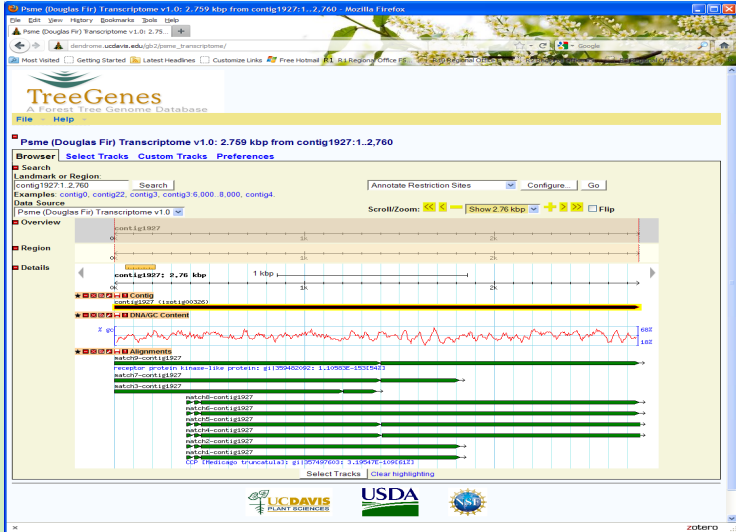# Climate Change Transcriptome Observatory
Things to come

- Douglas-fir annual atlas of transcript dynamics.
  - Relate transcriptome variation to phenology
  - Characterize expression from trees that are cold, mesic and warm adapted
  - Build networks of gene interactions to identify higher order responses
- Diurnal atlas of the Douglas-fir needle transcriptome.

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Douglas-fir transcriptome at Dendrome

http://dendrome.ucdavis.edu/gb2/psme_transcriptome/

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Dendrome annotation
52,309 total transcripts

33,145	blastx match (NCBI nr)
6,428	blastn to *Picea sitchensis*
12,736	no description
52,309	Total (clustered) transcripts

Thank you Jill Wegrzyn and Pine RefSeq project!

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

Acknowledgements

The Douglas-fir needle transcriptome.
Differential expression: the dormancy-spring growth transition.
Annotation: are differentially expressed isogroups annotated?

391 differentially expressed isogroups
111 blastx hits (96 hits to plant, 13 hits to bacteria/fungi)

- Cell wall/glucan processes
  - Xyloglucosyl transferase
  - Polygalacturonase
- Energy production
  - Proline-rich APG-like protein
  - PEP Carboxykinase
  - Glyceraldehyde-3-phosphate dehydrogenase
- Carbohydrate/Metabolism
  - Beta-glucosidase

- Transcription factors
  - NAM (no apical meristem)-like
  - WRKY11 transcription factor
  - AP2/EREBP transcription factors
  - F-box/KELCH
  - ETHYLENE responsive elements binding factors
  - Protease-associated zinc finger
  - Protein phosphatase 2C(maintanence of meristem activity)
  - R2R3 MYBs Myb2, myb 305
  - Transducin wd40-domain