

Large-scale RNA-Seq Transcriptomics Studies in *P. tremula*: Projects and Resource Development

Nathaniel Street



Umeå Plant Science Centre
a centre of excellence

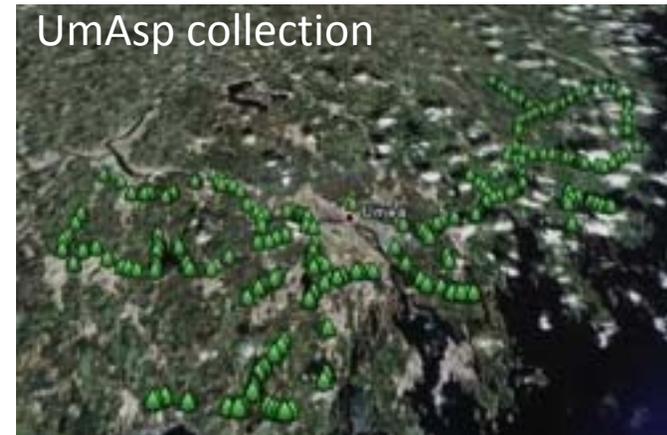


The UmAsp Sex Project

Agilent arrays: 10 male, 10 female (2008)

RNA-Seq: 8 male, 8 female + 4 poplars (2008+2010)

Phenotypic data: leaf traits + arthropods (2010)



Signal or Noise?

gw1.III.2363.1

- V1.0 – present
- V1.1 – absent but region in assembly
- V2.x – region absent in assembly

Multiple analysis methods

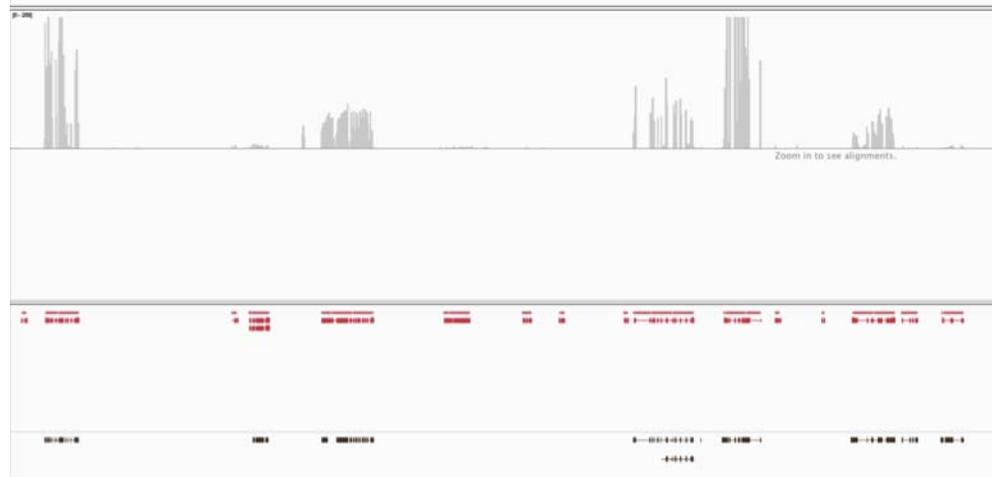
- Univariate
- Machine learning**



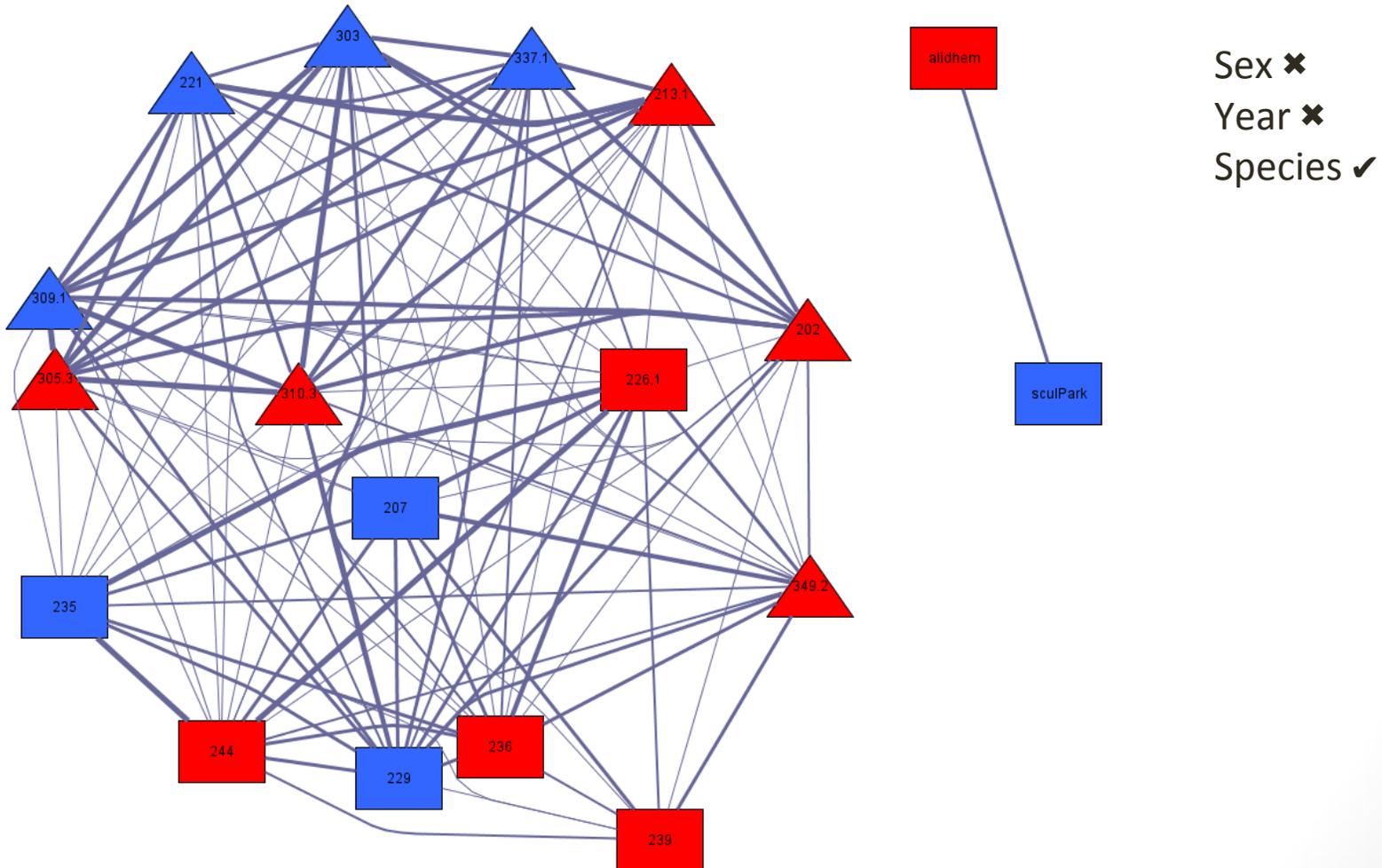
➔ RNA-Seq

Read mapping

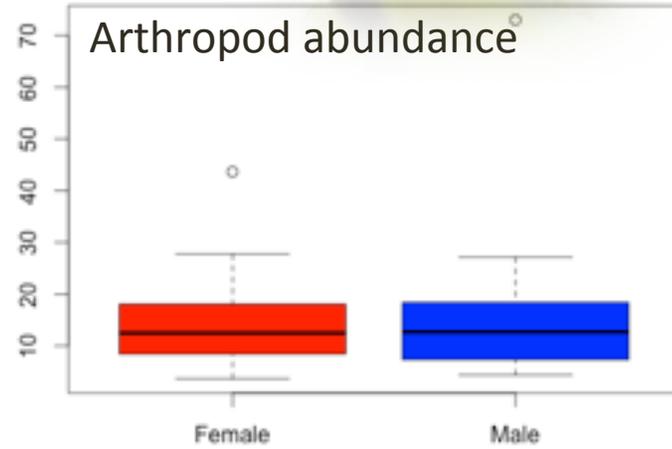
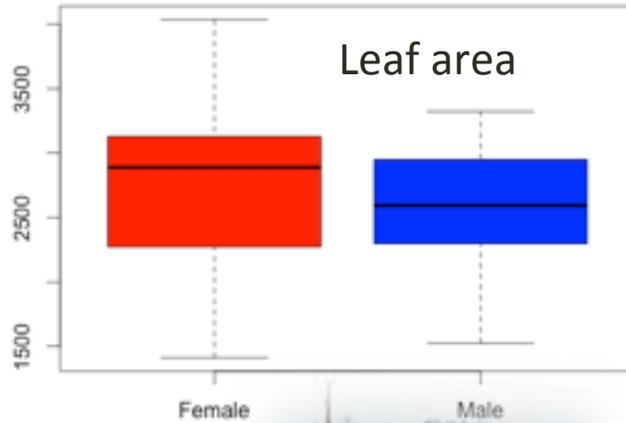
- Mapping to *P. trichocarpa* transcriptome
 - Cufflinks alignment to genome results in unstable expression profiles
 - RSEM much more reliable
- Raw data:
 - ~25% map to *P. trichocarpa*
 - ~50% map to *P. tremula*
- With QC trimming:
 - Over 80% map to *P. tremula*
 - ~50% map to *P. trichocarpa*



RNA-Seq Sample Network

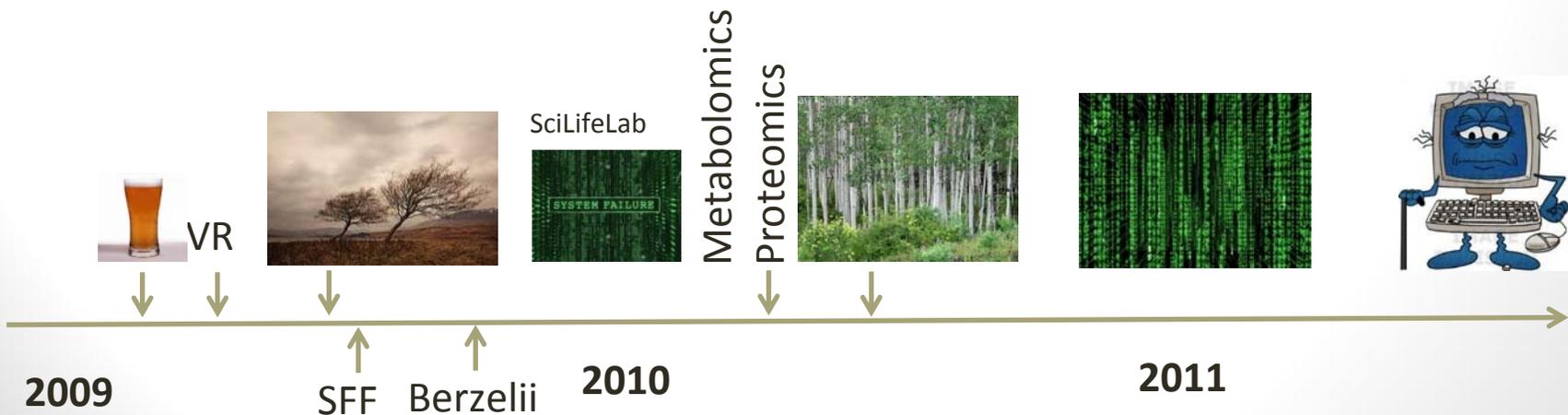


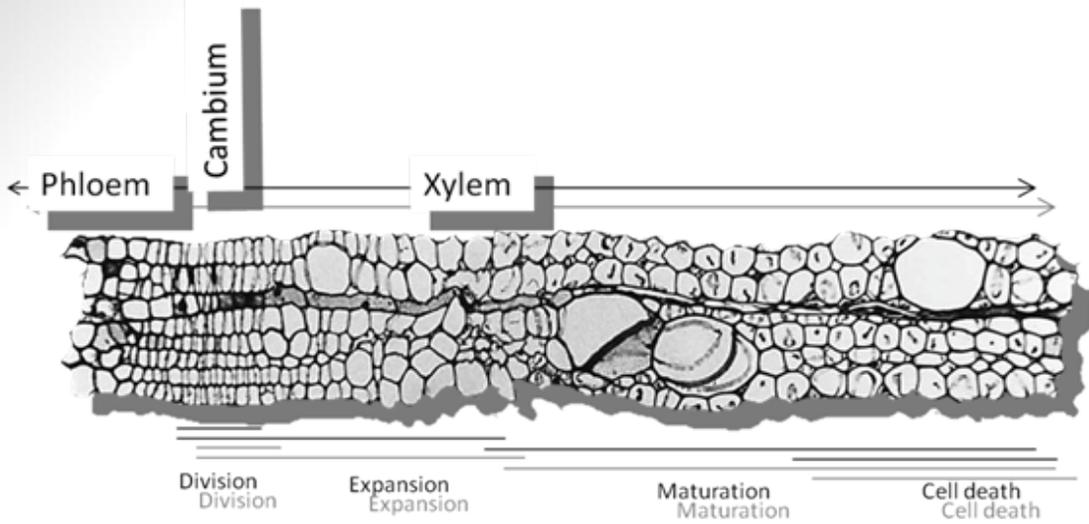
Sexual Equality



The AspWood Project

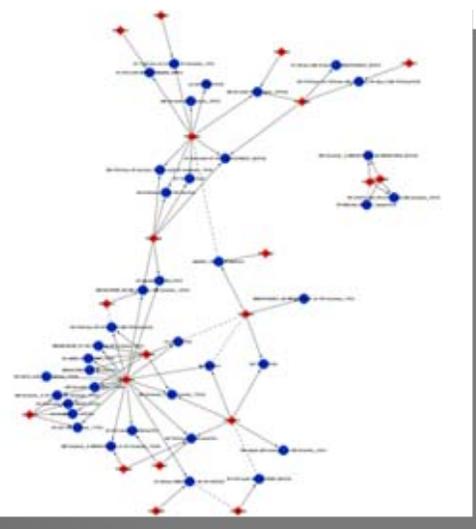
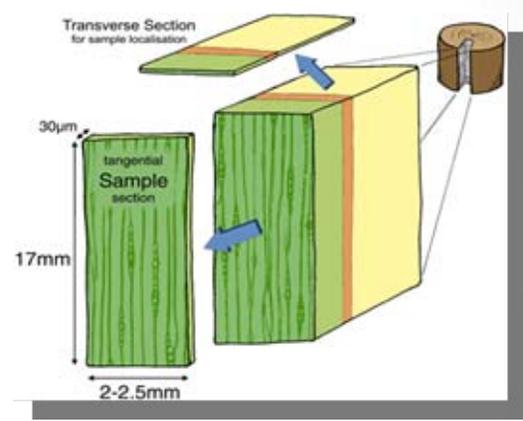
- 14 group leaders involved:
 - Catherine Bellini, Rishi Bhalerao, Stefan Jansson, Jan Karlsson, Ewa Mellerowicz, Thomas Moritz, Totte Niittyta, Ove Nilsson, Johan Trygg, Hannele Tuominen; Gunnar Wingsle, Nathaniel Street, Björn Sundberg, **Torgeir Hvidsten**
- Financial support:
 - The SSF centre for Developmental Biology of Plants
 - UPSC Berzelii Centre for Forest Biotechnology



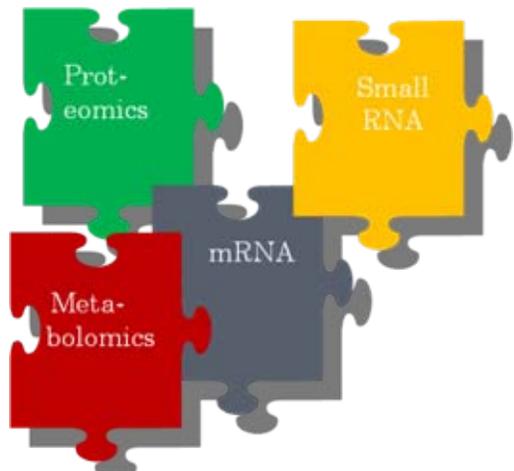


Sampling:

- 30 sections from phloem to cell death
- five trees



Reverse engineering



RNA-Seq
GC- and LC-MS

The trees

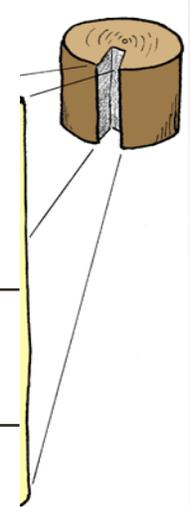
- July 7th, 2010 at 10.00-12.00. Vindeln/Mullkälen.
- 10 trees, 15 m high, about 47 years old.
- Leaves were also sampled for genotyping: genetically identical
- Propagated from roots for leaf developmental series (2011)



Samples

- Tangential cryo-sections of 20 μm sampled from phloem to cell death
- Sections for mRNA sequencing are pooled into 30 samples
- Sections for small RNA sequencing are pooled into 5 samples
- Min 20 M mRNA RNA-Seq read pairs
- Current analysis based on alignment to *P. trichocarpa* transcriptome

	Sections #	Sample #	Sample Name	sRNA
	1--3	3	K9-00	Old Phloem
	4--8	5	K9-01	Phloem
	9--9	1	K9-02	Cambium1
	10--10	1	K9-03	Cambium2
	11--11	1	K9-04	Cambium3
	12--12	1	K9-05	Cambium4
	13--13	1	K9-06	Cambium5
	14--14	1	K9-07	Cambium6
	15--15	1	K9-08	Cambium7
	16--16	1	K9-09	Cambium8
	17--19	3	K9-10	Exp/DevXyl1
	20--22	3	K9-11	Exp/DevXyl2
	23--25	3	K9-12	Exp/DevXyl3
17n	26--28	3	K9-13	Exp/DevXyl4
	29--31	3	K9-14	Exp/DevXyl5
	32--34	3	K9-15	Exp/DevXyl6
	35--37	3	K9-16	Exp/DevXyl7
	38--40	3	K9-17	Exp/DevXyl8
	41--43	3	K9-18	Exp/DevXyl9
	44--46	3	K9-19	Exp/DevXyl10
	47--55	9	K9-20	MatXyl1
	56--64	9	K9-21	MatXyl2
	65--73	9	K9-22	MatXyl3
	74--82	9	K9-23	MatXyl4
	83--91	9	K9-24	MatXyl5
	92--100	9	K9-25	MatXyl6
	101--109	9	K9-26	MatXyl7
	110--118	9	K9-27	MatXyl8
	119--127	9	K9-28	MatXyl9
	128--135	8	K9-29	MatXyl10



g PNAS 19

Expression consistency

Correlations for gene POPTR_0004s23690

K2 K3 0.97

K1 K9 0.96

K1 K2 0.96

K3 K9 0.96

K2 K9 0.95

K1 K3 0.95

K1 K5 0.88

K3 K5 0.88

K2 K5 0.85

K5 K9 0.82

Median > 0.7

With at least one mapped read:

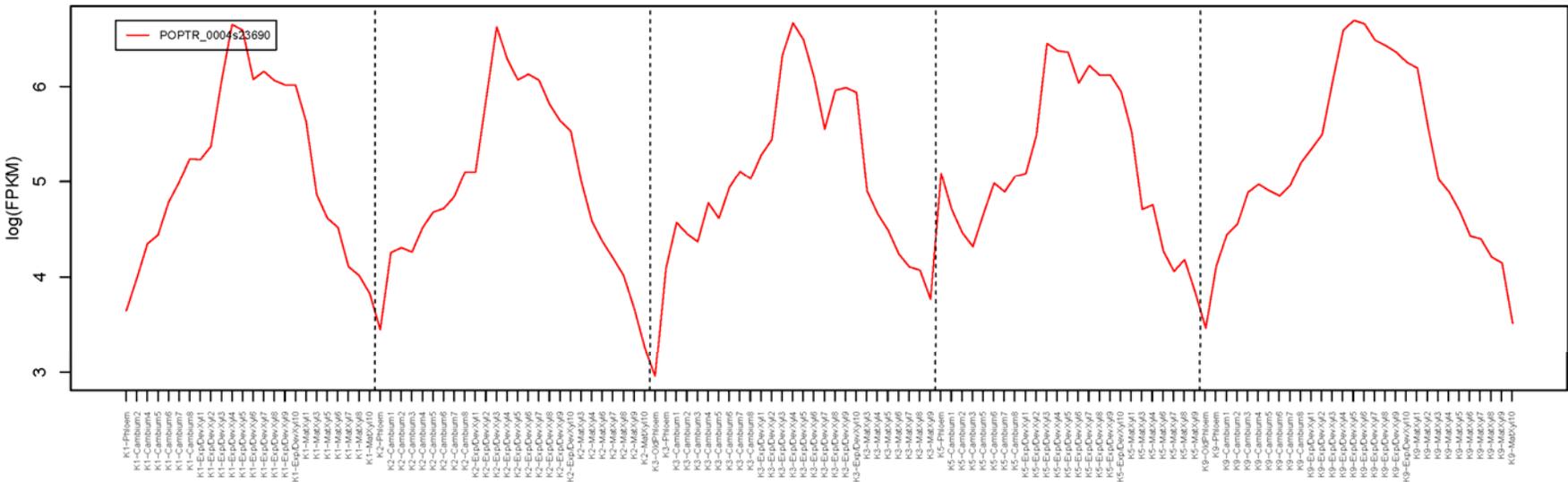
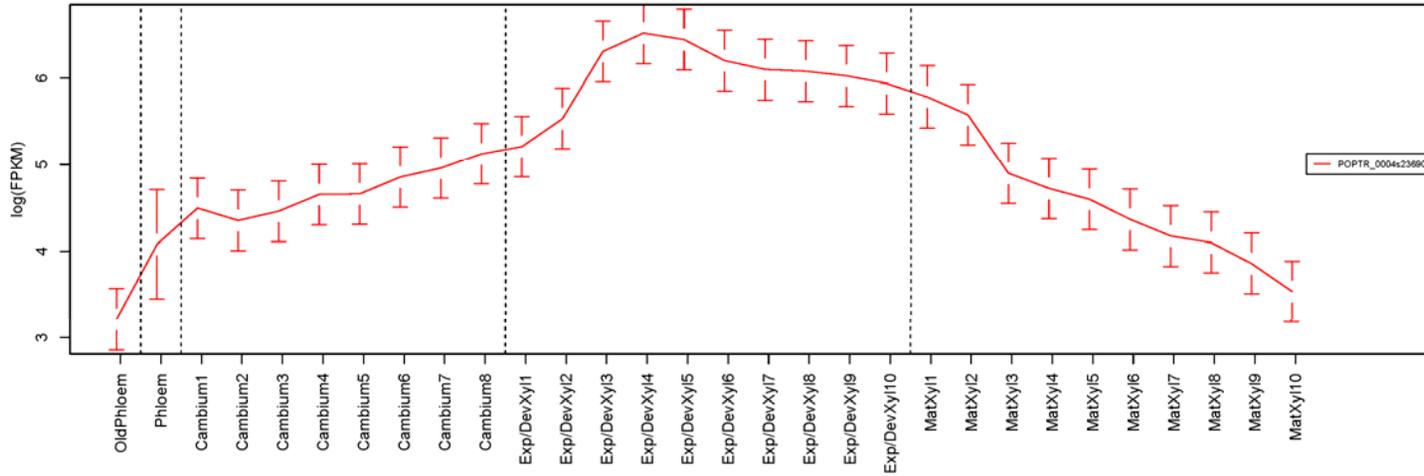
- Genes: 29148
- Transcription factors: 1467

With consistent expression and std. dev > 0.75:

- Genes: 8747
- Transcription factors: 521

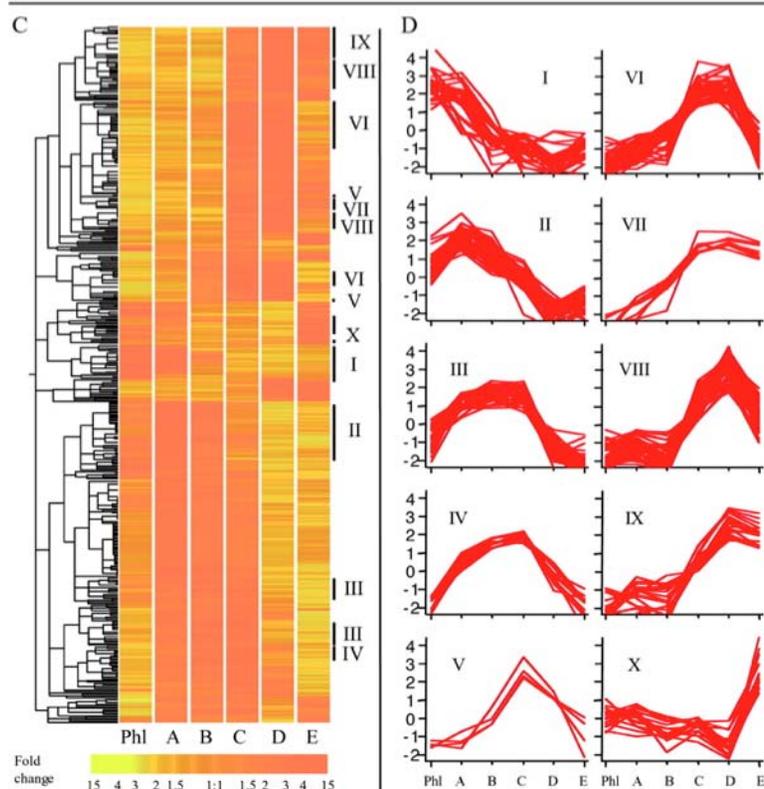
Example:

POPTR_0004s23690, Median= 0.95

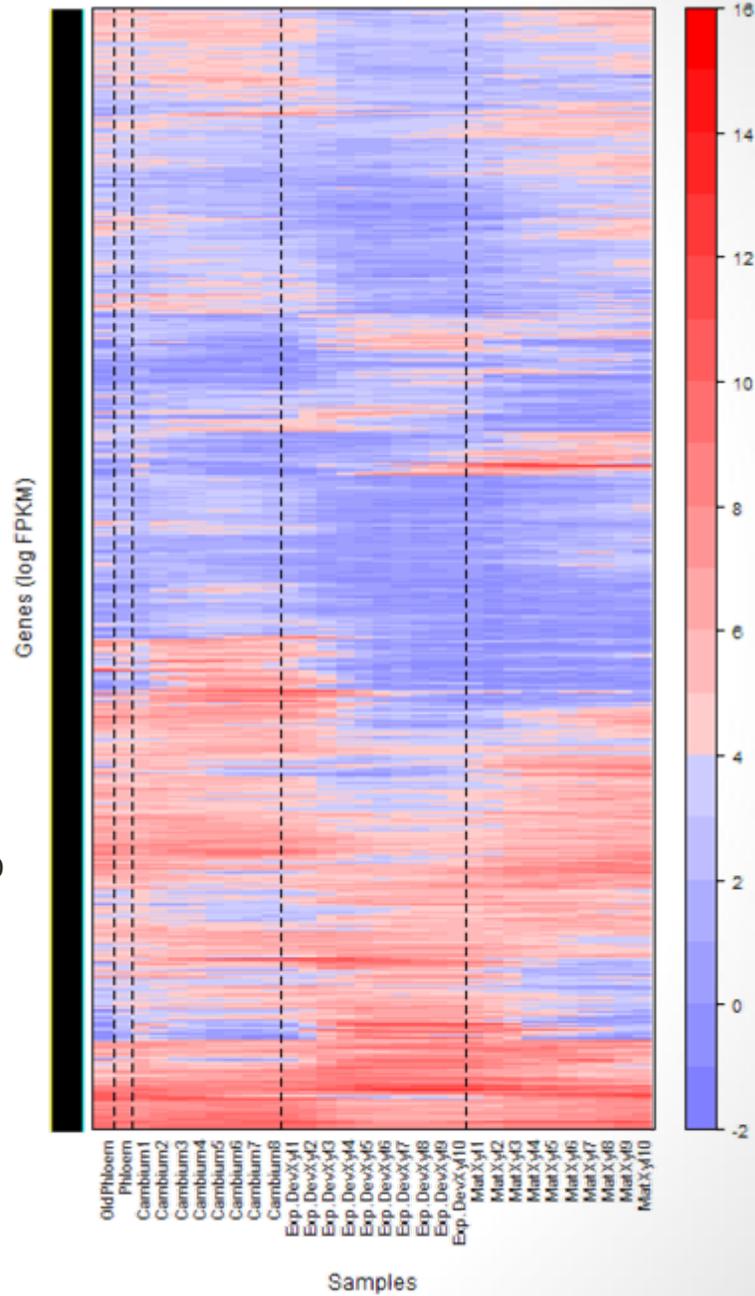


mRNA: Genes

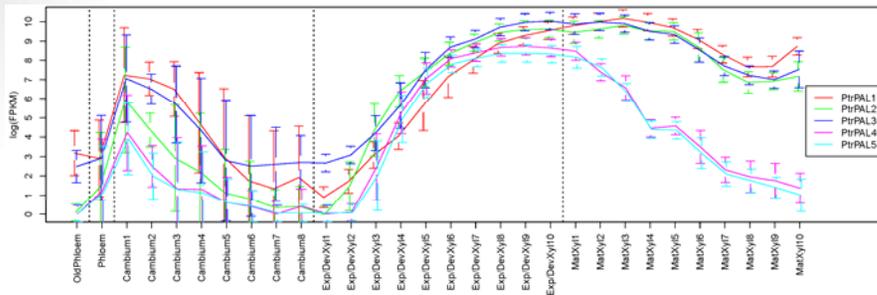
Hertzberg et al, (2001), PNAS, 98, 14732-14737



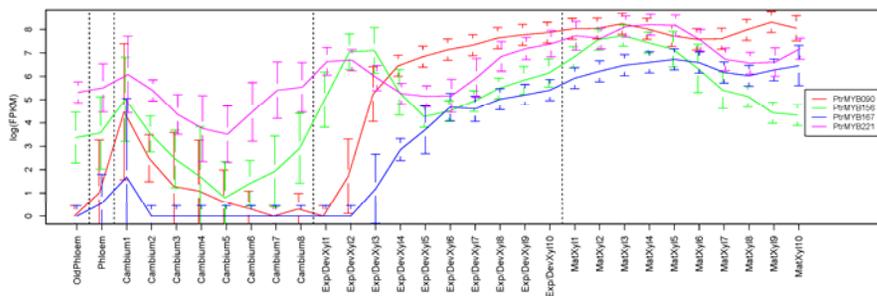
1791 selected genes



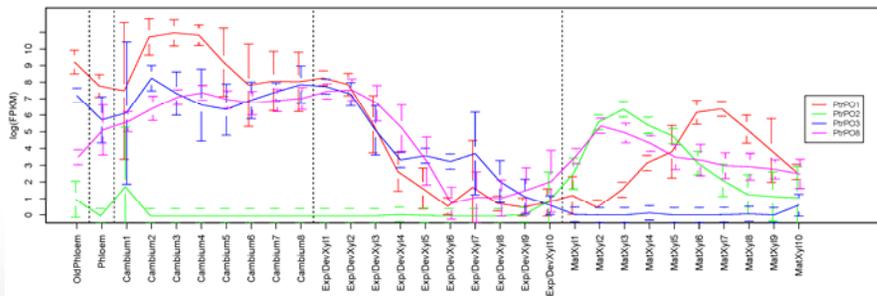
Phenylalanine ammonia-lyase



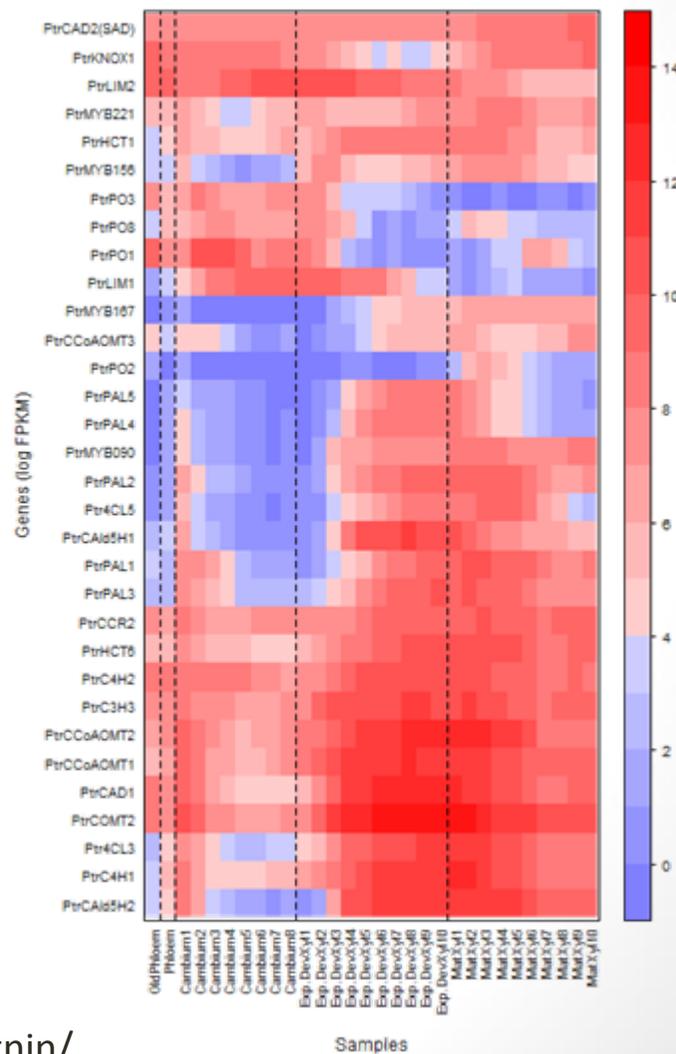
MYB



Peroxidase

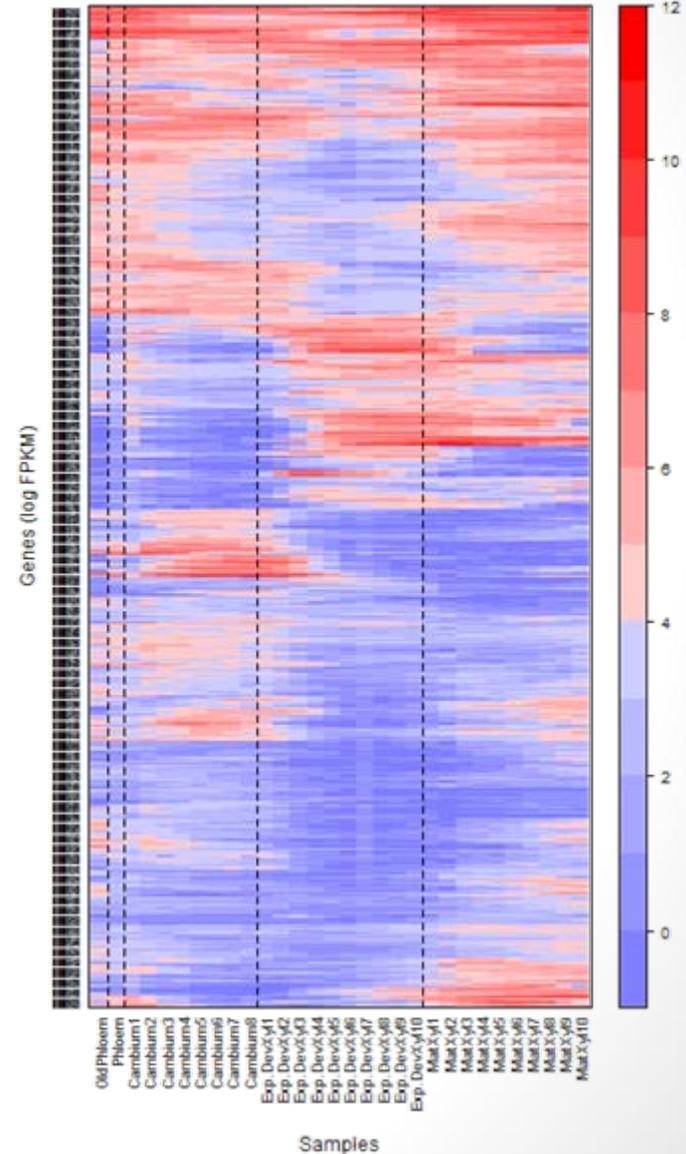
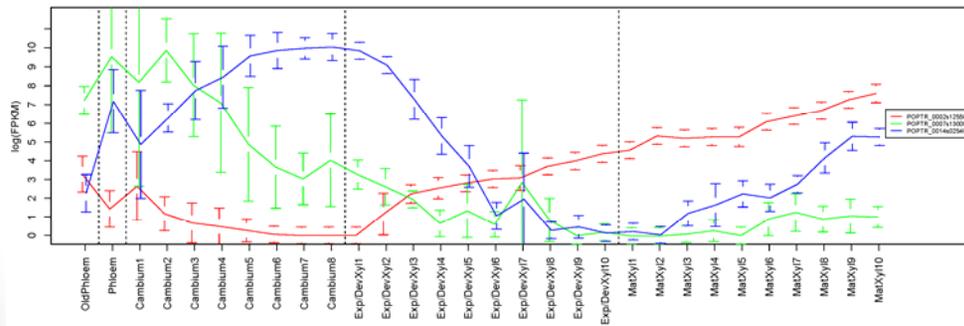
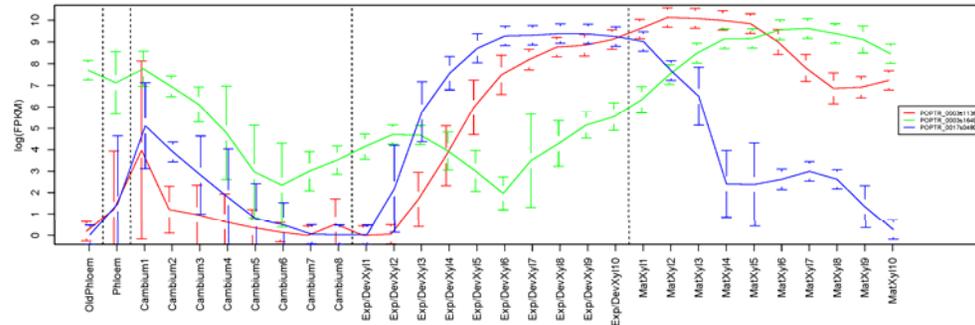


32 lignin genes



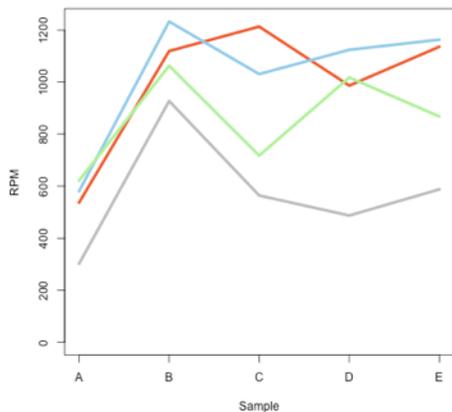
mRNA: Transcription factors

521 transcription factors

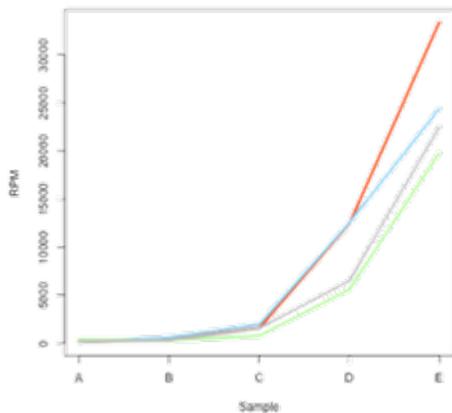


miRNA

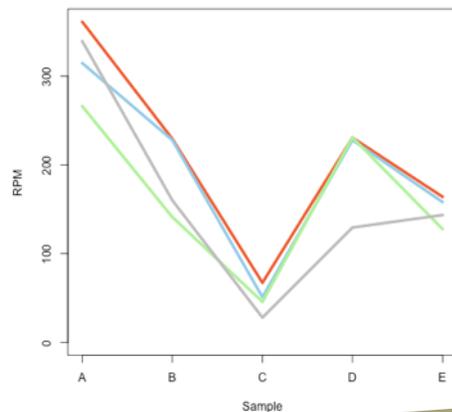
MIR481



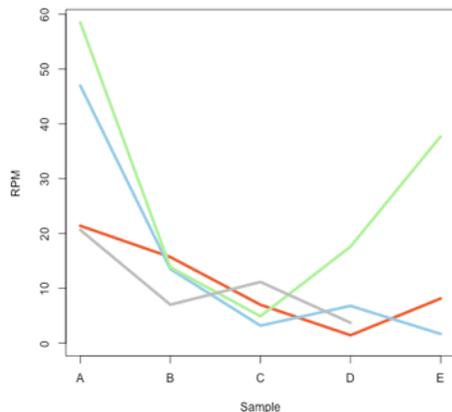
MIR390



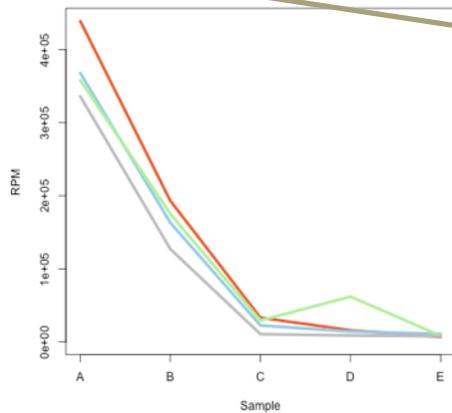
MIR164



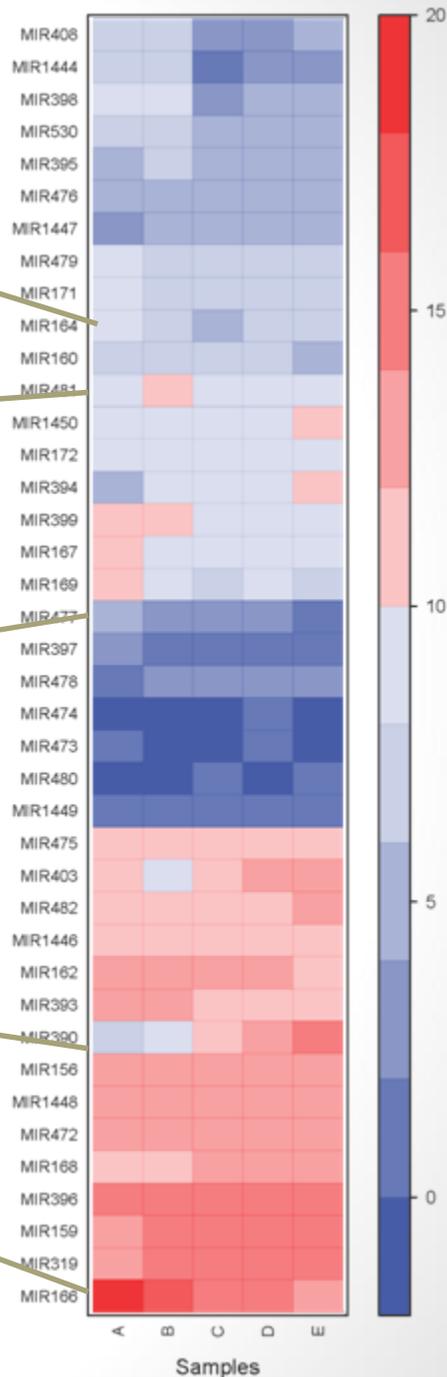
MIR477



MIR166

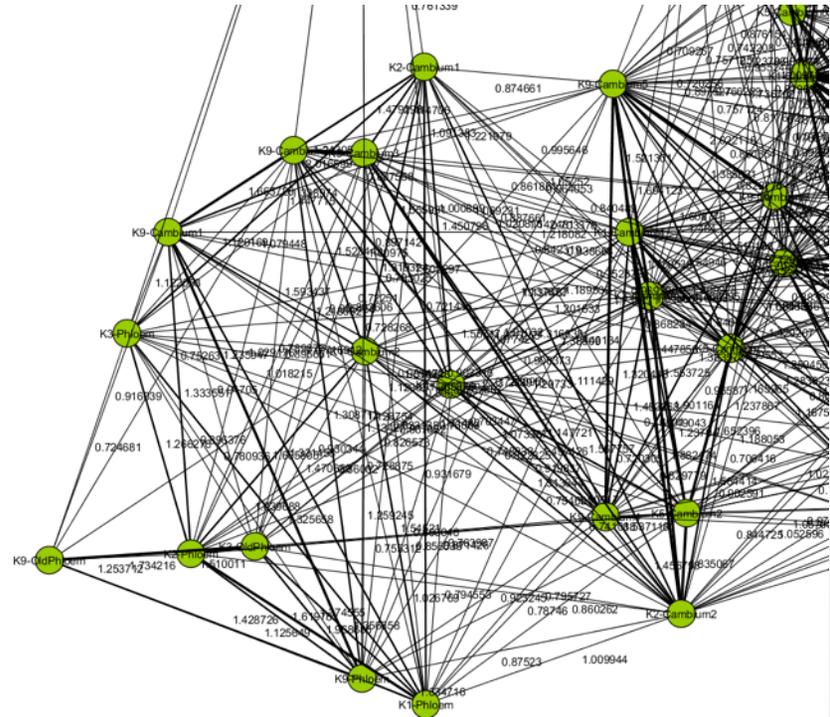
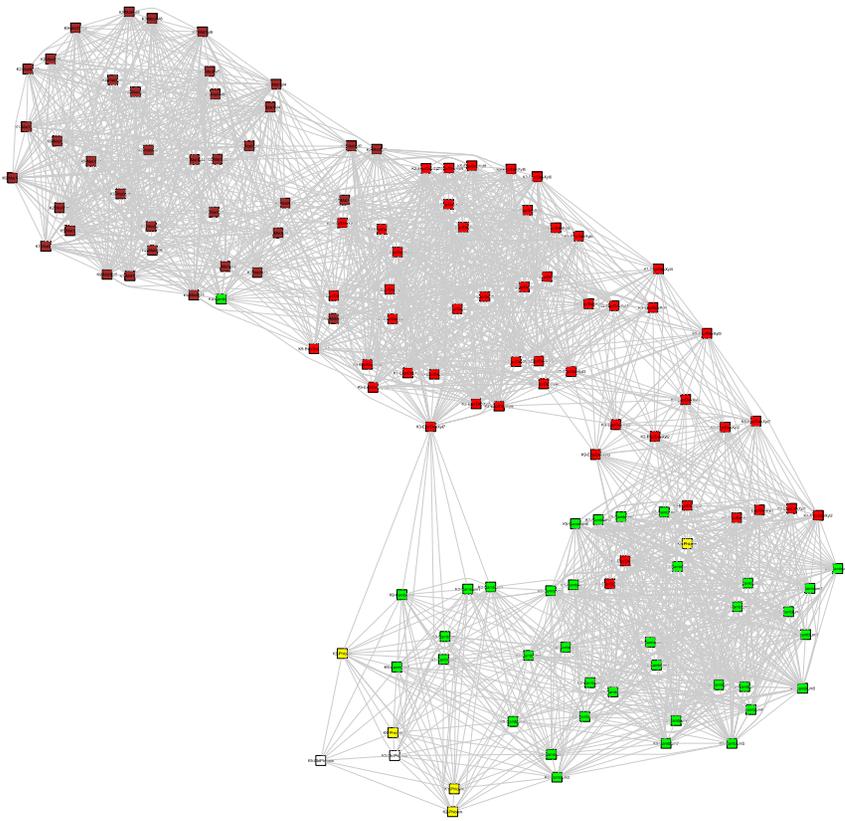


miRNA families (log RPM)

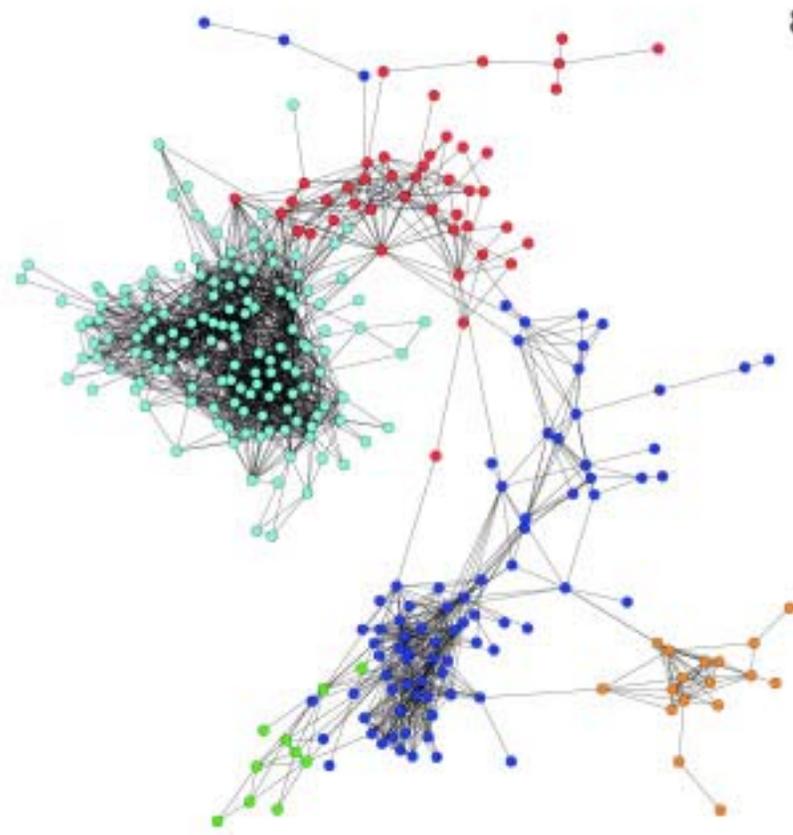
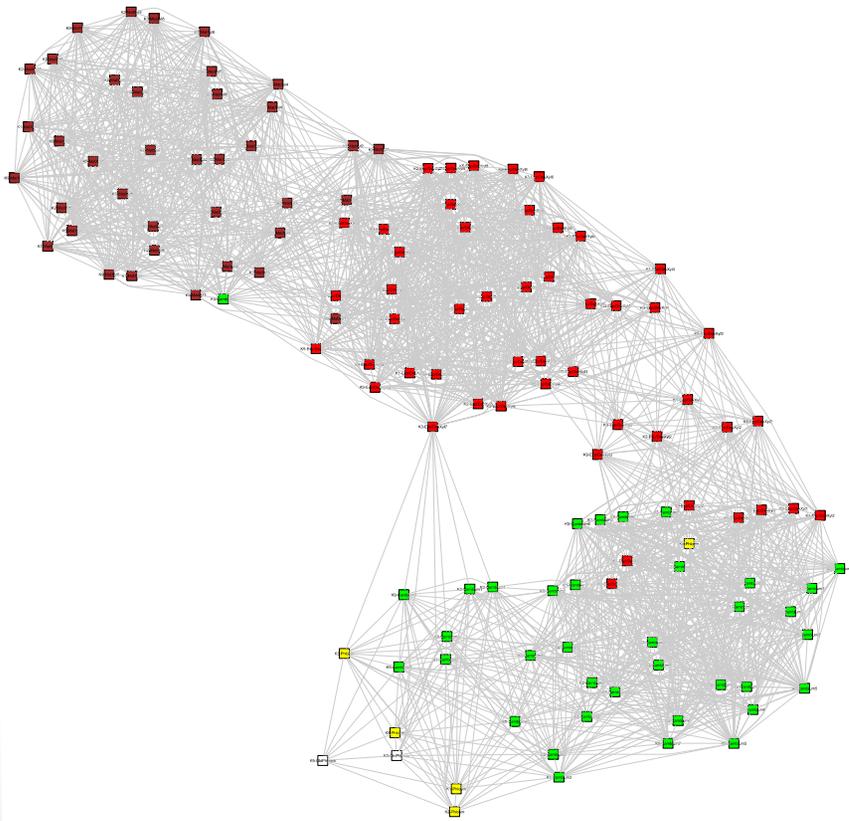


- A Phloem
- B Cambium
- C Exp/Dev Xyl
- D Exp/Dev Xyl
- E Mat Xyl

Network Analysis

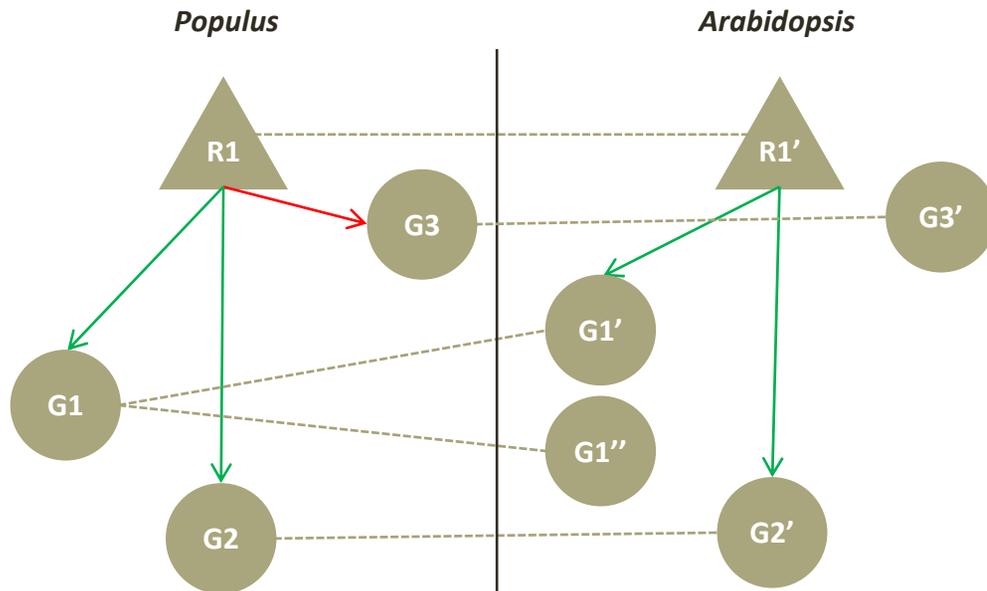


Network Patterns



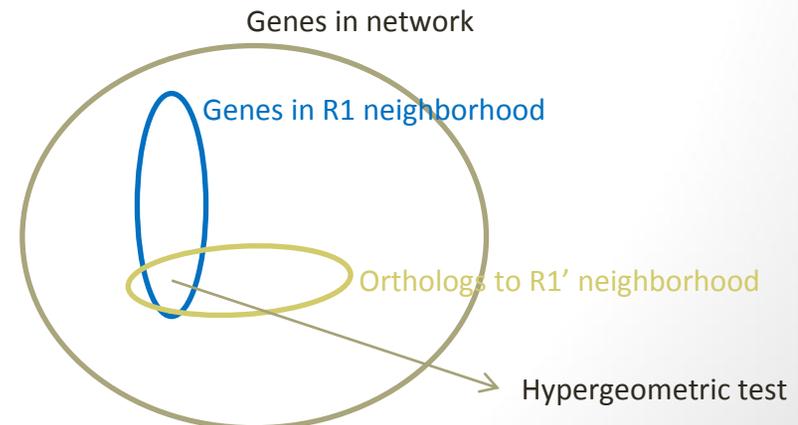
Street *et al.* 2008

Comparative regulomics



Identify conserved regulatory control

- Public data (arabidopsis, rice, poplar)
- **Spruce Genome Project**
- UPSC aspen projects



Spruce Genome Project

- Phase two projects
 - Wood development series (90)
 - Compression wood formation (90)
 - Diurnal expression (72)
 - Seedling development and megagametophyte (30)
 - Somatic culture → differentiation series (50)
 - Tracheary element differentiation (54)



PopGenIE – Case study

- All UPSC RNA-Seq data will be made available and integrated into PopGenIE (and spruce into ConGenIE)
- Public Galaxy server
- Community based functional annotation
- Literature incorporation
- Example use:
 - POPTR_0019s13500: secretory carrier-associated membrane protein
 - Part of the BioImprove collection I



eFP Browser | Popgenie v2.0 x

popgenie.org/tool/efp-browser

Home

eFP Browser

Selection -1,0

Tools

- ePlant-tissue expression
- eFP-tissue expression
- PopNet - network layout
- ExPlot - gene profiles
- ExCluster - heatmap clustering
- GO enrichment
- GO-Tree
- Chromosome Diagram

Data Source: Tissue comparison

Gene model name: POPTR_0019s13500

Signal Threshold: 1.58

- Selected Genes

POPTR_0019s13500

POPTR_0019s13500
protein_coding

UPSC-Tissue Comparison
UPSC-BASE: UMA-0030
Sjödin et al. 2008

Flower catkin

Leaf 8

Leaf 10

Old root

Apical region

Bark

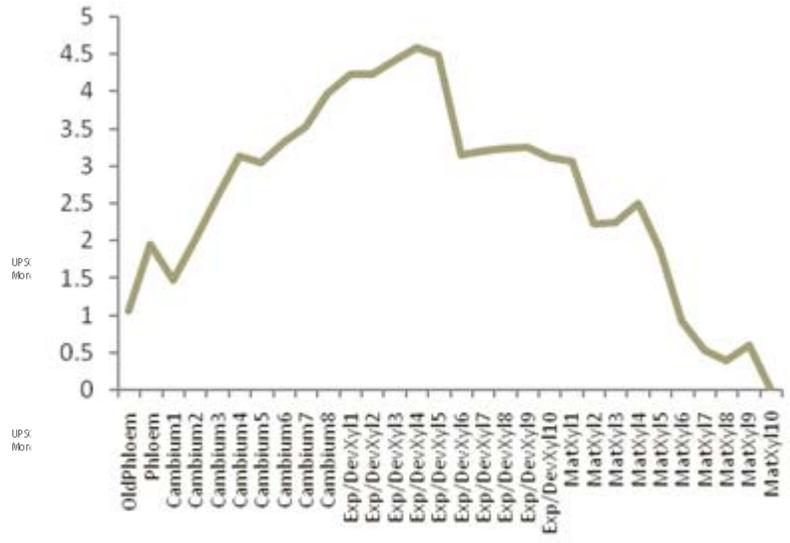
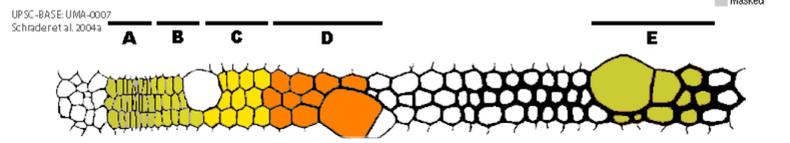
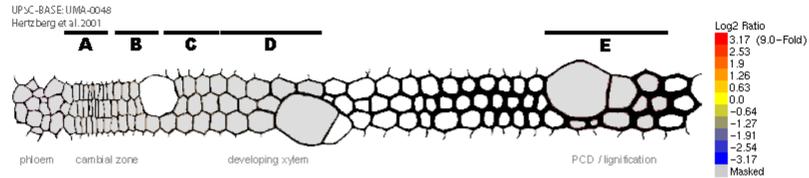
Phloem

Xylem

Log2 Ratio (2.9-Fold)

1.58
1.26
0.94
0.63
0.31
-0.01
-0.32
-0.64
-0.95
-1.27
-1.58
Masked

- eFP-tissue expression
- PopNet - network layout
- ExPlot - gene profiles
- ExCluster - heatmap clustering
- GO enrichment
- GO-Tree
- Chromosome Diagram



phloem cambium fiber extension secondary cell wall PCD / lignification

popNet | Popgenie v2.0

popgenie.org/tool/popnet

popgenie.ORG

Home Info Tools About us Community Annotation Help Login

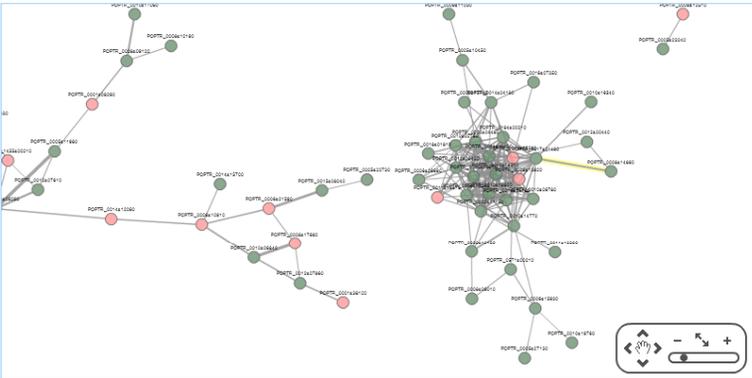
Home

popNet

Selection -301,0

Tools

- ePlant - tissue expression
- eFP - tissue expression
- PopNet - network layout
- ExPlot - gene profiles
- ExCluster - heatmap clustering
- GO enrichment
- GO-Tree
- Chromosome Diagram



GO selection:

- Biological Process
- Cellular Component
- Molecular Function

There are no GOs selected you can select them [here](#)

Pink nodes are genes in your selected GO

When expanding your network red genes are in your GO selection

Network: CLR

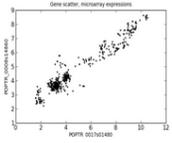
Threshold: 7 Selection: Add Remove

ExpandTh: 4

Correlation: CLR

Nodeshape: Node New selection: Select

Geneprofiles: Static Network(NA) Display



geneprofile

scatterplot

heatmap

Info

The scatterplot shows expression values of the first selected gene against other genes. (limit 2-7 genes)

GO enrichment | Popgenie

popgenie.org/content/go-enrichment

Home

GO enrichment

Selection - 301,0

Tools

- ePlant-tissue expression
- eFP-tissue expression
- PopNet - network layout
- ExPlot - gene profiles
- ExCluster - heatmap clustering
- GO enrichment
- GO-Tree
- Chromosome Diagram

Properties Window

GO:0007018
microtubule-based movement

GO:0007018
microtubule-based movement
info: 9/122 | 82/10632
p-val: 3.18E-7

GO:0006909 phagocytosis

GO:0010324 membrane invagination

GO:0007017 microtubule-based process

GO:0071841 cellular component organization of biological cell

GO:0010044 cellular membrane organization

GO:0008211 cell death

GO:0071842 cellular component organization

GO:0071840 cellular organization

GO:0081024 membrane organization

GO:0012201 programmed cell death

Tree type:

Orientation:

Static network:

Info

Select genes in the network by click or drag selection.
Hold shift key to extend your selection.
Right-click menu:

- Expand your network for the selected genes
- Select neighbors to view visible connections to the selected node
- Delete nodes in the visible network (obs! They will not be removed from your selection of genes)

10e-10 10e-9 10e-8 10e-7 10e-6 10e-5 10e-4 10e-3 10e-2 not significant 10e-2 10e-3 10e-4 10e-5 10e-6 10e-7 10e-8 10e-9 10e-10

Over representation p-value Under representation p-value

Firefox Experiment search | Popgenie v2.0
 popgenie.org/content/experiment-search?id=session

popgenie.ORG
 Home Info Tools About us Community Annotation Help Login

Home

Experiment search

Selection -301,10
 Gene -301
 Experiment -10
 Great, You have selected 10 samples and 301 genes. Now you can use following tools.
 Tools

Collapse All	Sample ID	Experiment_info	Sample id
<input type="checkbox"/>	GSE16459	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	wood
<input type="checkbox"/>	GSM413780	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413782	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413781	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413778	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413777	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413776	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413783	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413784	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	
<input type="checkbox"/>	GSM413779	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	Wood formation;
<input type="checkbox"/>	GSM413785	circad_pop1-circadian rhythm in the expression of genes involved in the formation of poplar wood;	Wood formation;

Microarray ID: 0
 Sample id Search
 Add to Basket
 Remove from Basket
 Export table to CSV
 Show All Experiments

Number of samples: 10
 Transparent:
 Import

Experiment_info Sample description Treatment level ENVO description ENVO_id
 Experiment_des PO description Treatment protocol Treatment type PO id

Firefox

Heatmap | Popgenie v2.0

popgenie.org/content/heatmap

populus Search

Home Info Tools About us Community Annotation Help Login

Home

Selection -301,10

Tools

- ePlant-tissue expression
- eFP-tissue expression
- PopNet - network layout
- ExPlot - gene profiles
- ExCluster - heatmap clustering
- GO enrichment
- GO-Tree
- Chromosome Diagram

Heatmap

Color Key and Density Plot

Density

Column Z-Score

Populus Heatmap

CSMAN137580
CSMAN137579
CSMAN137581
CSMAN137583
CSMAN137582
CSMAN137586
CSMAN137578
CSMAN137584
CSMAN137577
CSMAN137576

[Download table](#)

DistanceFunction: euclidean
Hierarchical clustering: ward
Genes not found:

Heatmap: Selection

Distance function

Hclust

Display

Site Map | © 2011 Popgenie.org. Site Powered by Drupal®.

Acknowledgements

- Torgeir Hvidsten
 - Sergiu Netotea
 - Jenny Önskog
- Björn Sundberg
- Stefan Jansson
- Benedicte Albrechtsen
 - Kathryn Robinson
- Pär Ingvarsson
 - Agneta Andersson
- SciLifeLab, Uppmax
- Spruce Genome Assembly Team



David Sundell
Chanaka Mannapperuma
Sergiu Netotea

