



PAG XX - Forest Tree
January 15, 2012



DIFFERENTIAL TRANSCRIPTOMES DURING FORMING OF EPIGENETIC MEMORY IN NORWAY SPRUCE

Igor Yakovlev

An epigenetic memory in *Picea abies*

It was established that temperature during embryo development in Norway spruce caused significant and long lasting phenotypic changes in the progeny
(Johnsen et al., 2005; Kvaalen & Johnsen, 2008)

Traits we know (so far) that are affected is the timing of

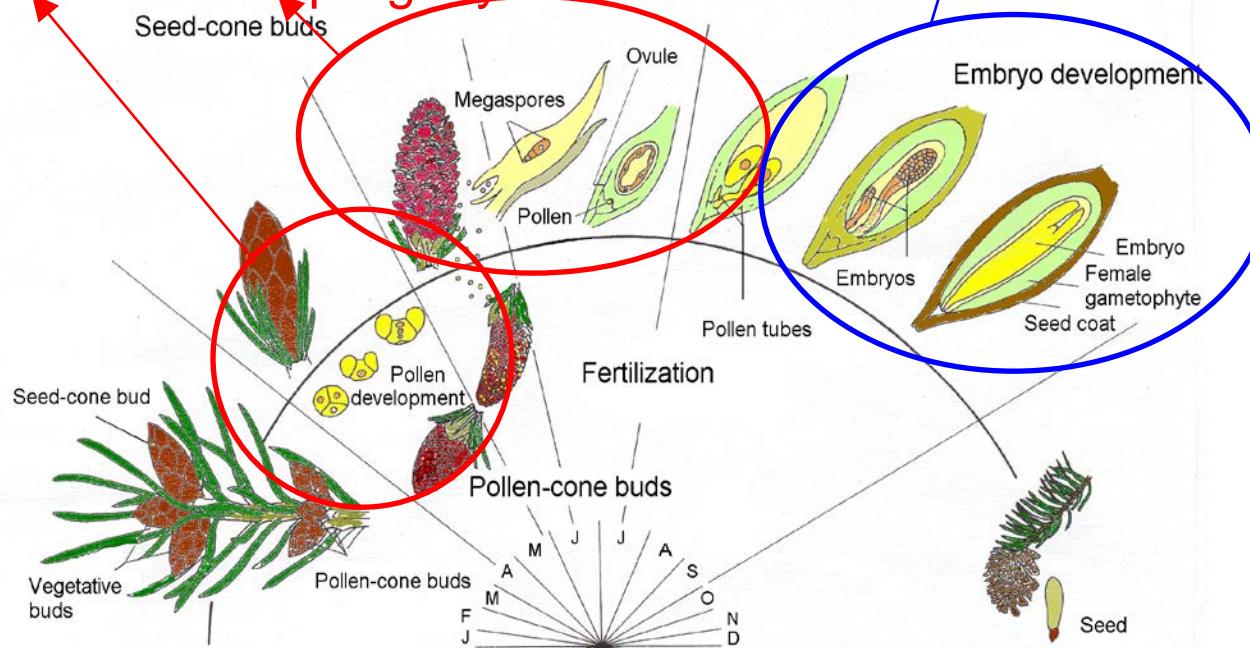
- dehardening and bud burst in the spring
- leader shoot growth cessation in the summer
- bud set and cold acclimation in the autumn

The timing **is delayed** if seeds are produced under warmer conditions than normal and **advanced** if seeds are produced **colder** than normal

Kvaalen & Johnsen, 2008; Scroppa at al., 2010

Timing of epigenetic memory mark establishment

Observed no effect of treatments on the progeny Phenotypic changes observed



Johnsen et al., 1996; 2005

Epigenetic memory marks establishment occur exclusively during embryogenesis in response to environmental impact, and the epitype is fixated by the time the embryo is fully developed without a change in the DNA sequence

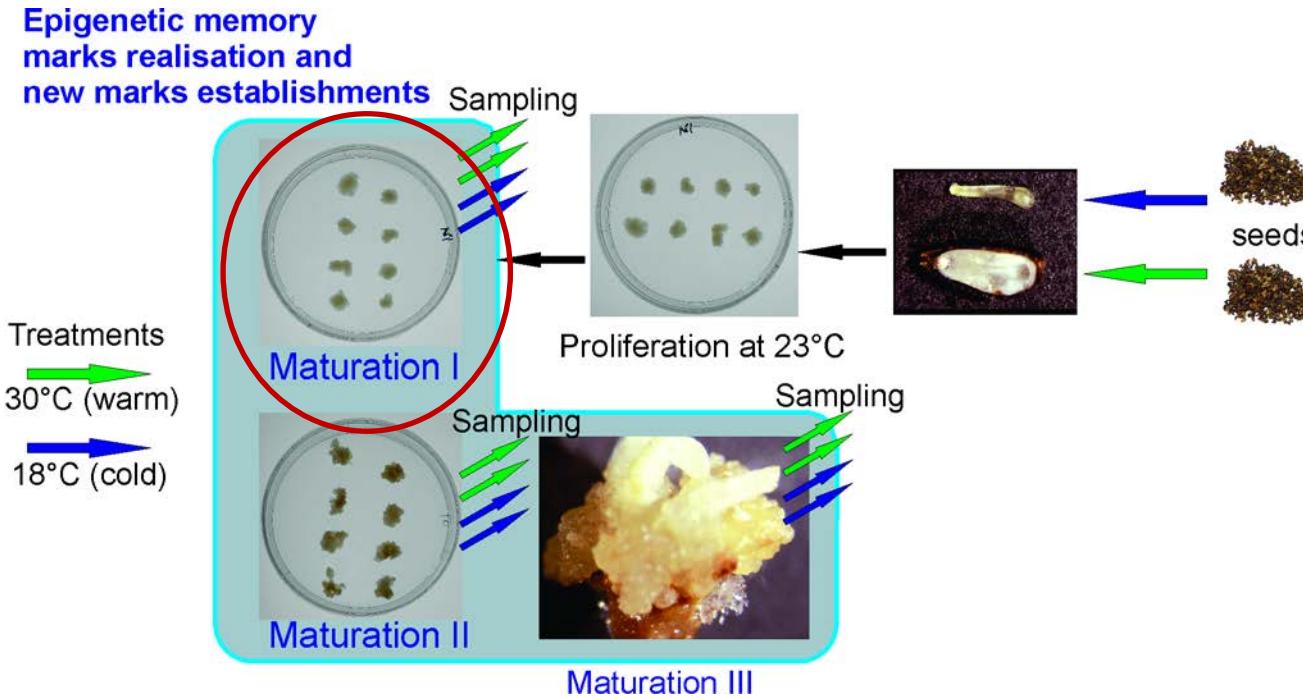
General experiment overview

We aimed on identifying and characterize genes and other regulatory elements involved in the initiation, maintenance, and heritability of epigenetic states using high throughput (next generation) sequencing approaches

Using high-throughput sequencing approaches: the 454 GS-FLX+ Titanium™ and Illumina GA II based MACE (massive analysis of cDNA ends) we analyzed *P. abies* transcriptomes (GenXPro)

General experiment overview

Samples

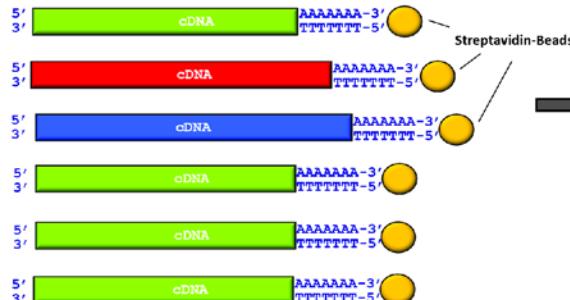


Zygotic embryogenesis cultures produced from embryos from two crosses of **cold** and **warm** originated “embryonic epitypes” growing at 18° C and 30° C at maturation I stage

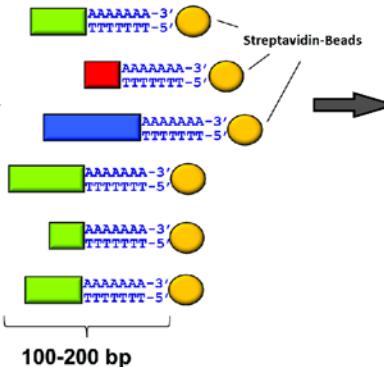
Description of MACE (massive analysis of cDNA ends)

GenXPro
Booth: 428

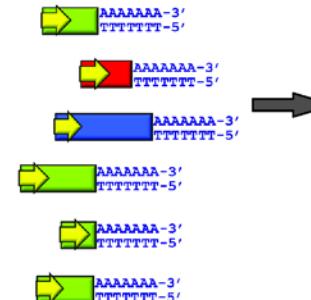
cDNA population bound to Streptavidine



Fragmentation



2nd generation sequencing of 50-100 bp

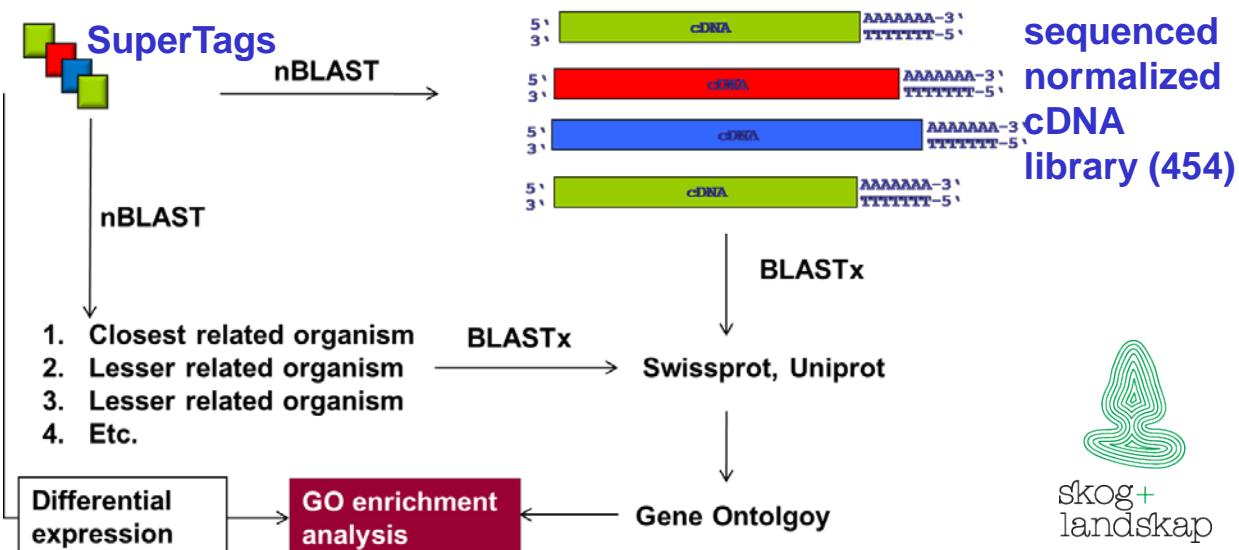


Assembly & Counting



- Longer tags are generated: up to 200 bp after assembly
- High depth of analysis: rare transcripts are included

Bioinformatics:
Functional
annotation



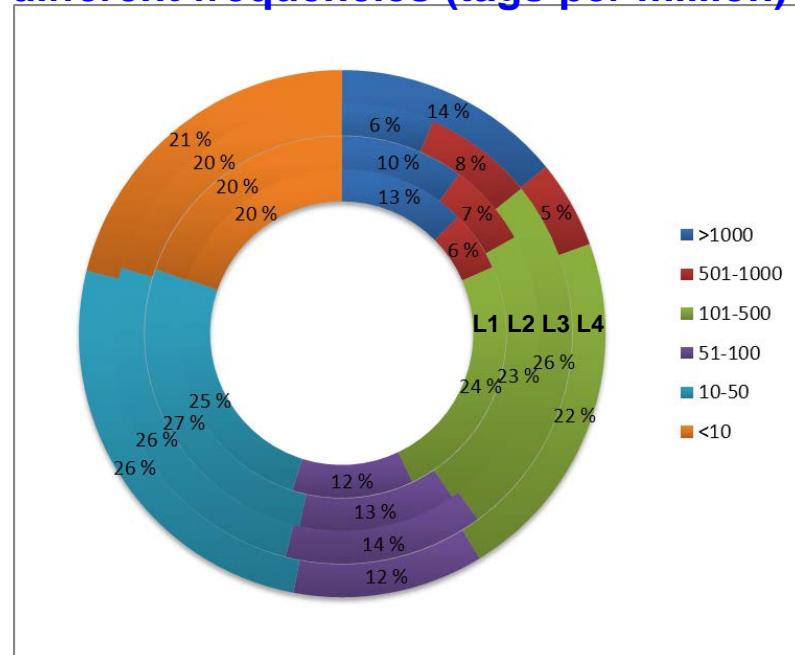
skog +
landskap

MACE results

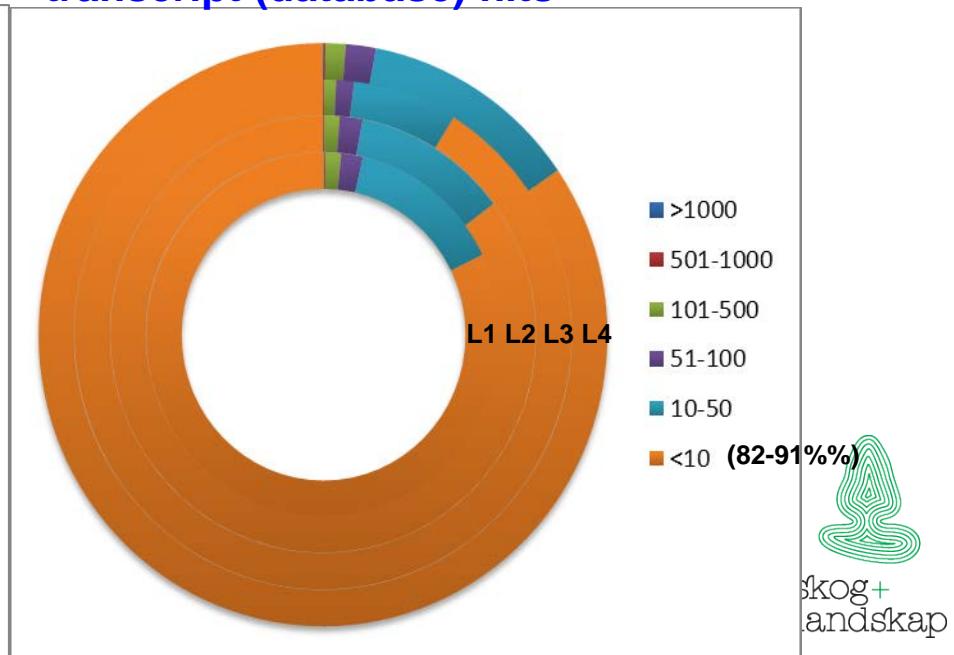
143 723 unique database (transcript) hits

Ver.3 (final)	Tags	Reads	L1 (W 18° C)	L2 (W 30° C)	L3 (C 18° C)	L4 (C 30° C)
Hits S+AS	2 645 680	5 678 042	472 487	902 217	3 706 088	597 250
no hit	2 291 996	2 855 900	299 977	471 234	1 715 947	368 742
Total (quality)		8 533 942	772 464	1 373 451	5 422 035	965 992
hit, %%			61,2	65,7	68,4	61,8
no hit, %%			38,8	34,3	31,6	38,2
unique transcripts			53 223	70 264	119 501	61 188
Max, tags per million			4 349,7	6 705,7	2 763,2	21 216,5

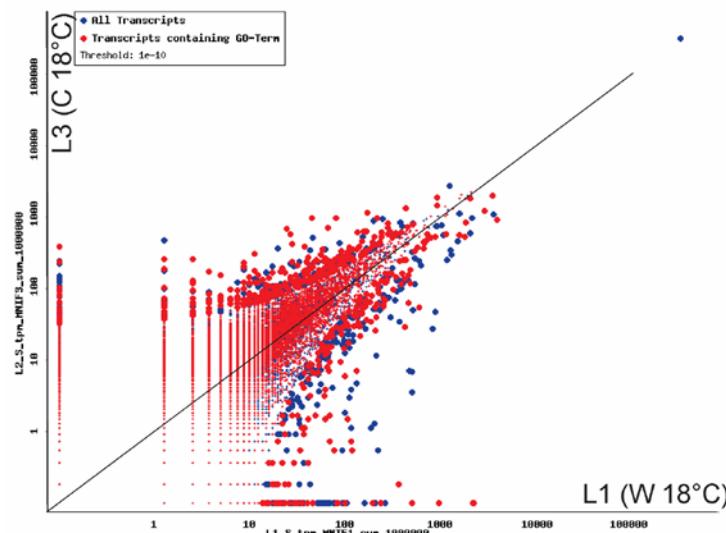
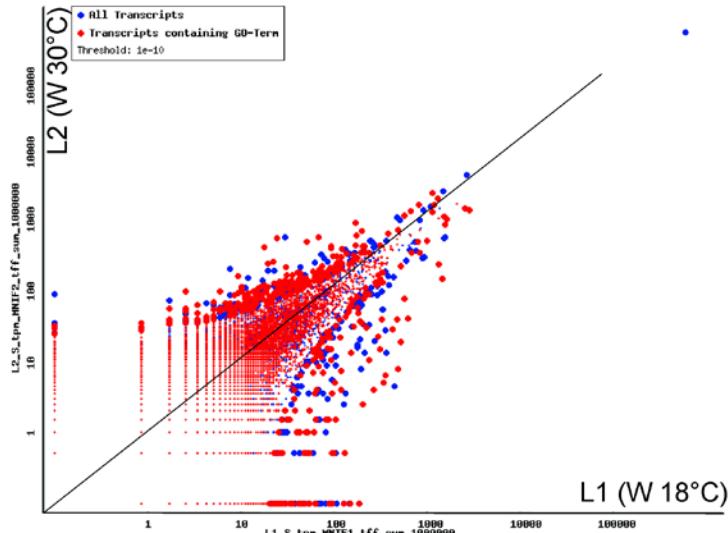
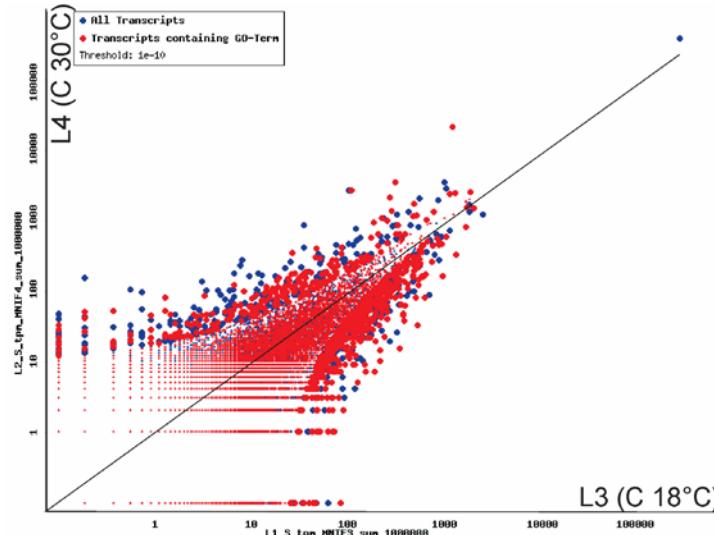
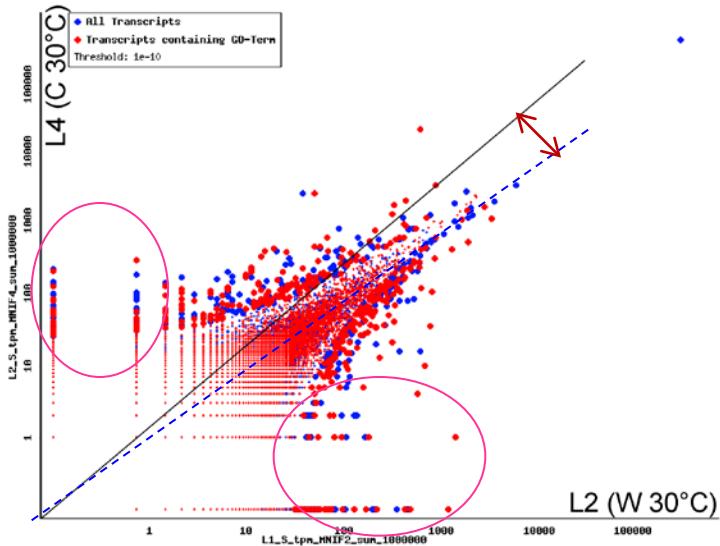
Distribution of unique transcripts with different frequencies (tags per million)



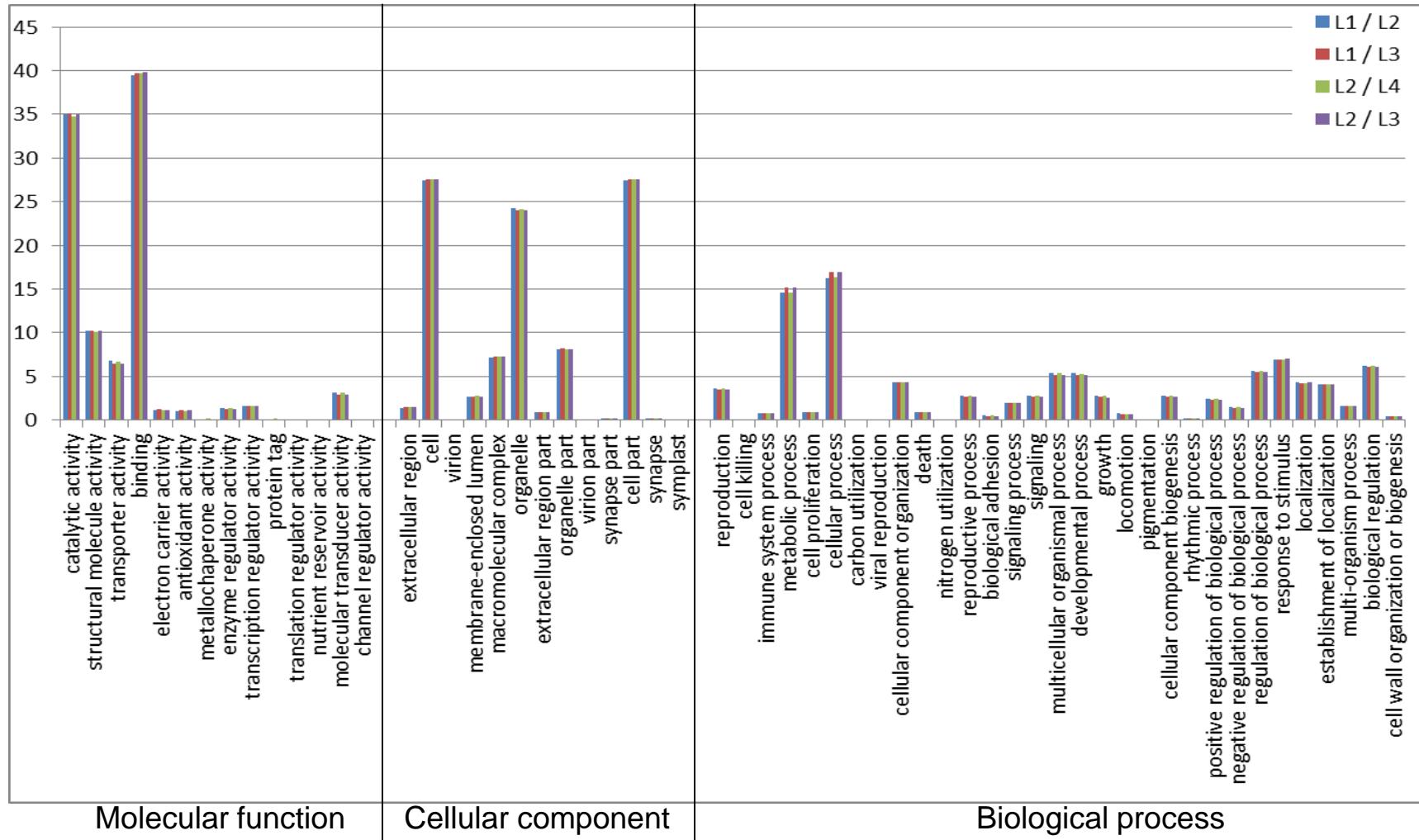
Distribution of total amount of different transcript (database) hits



Scatterplots for GO-category – cellular components

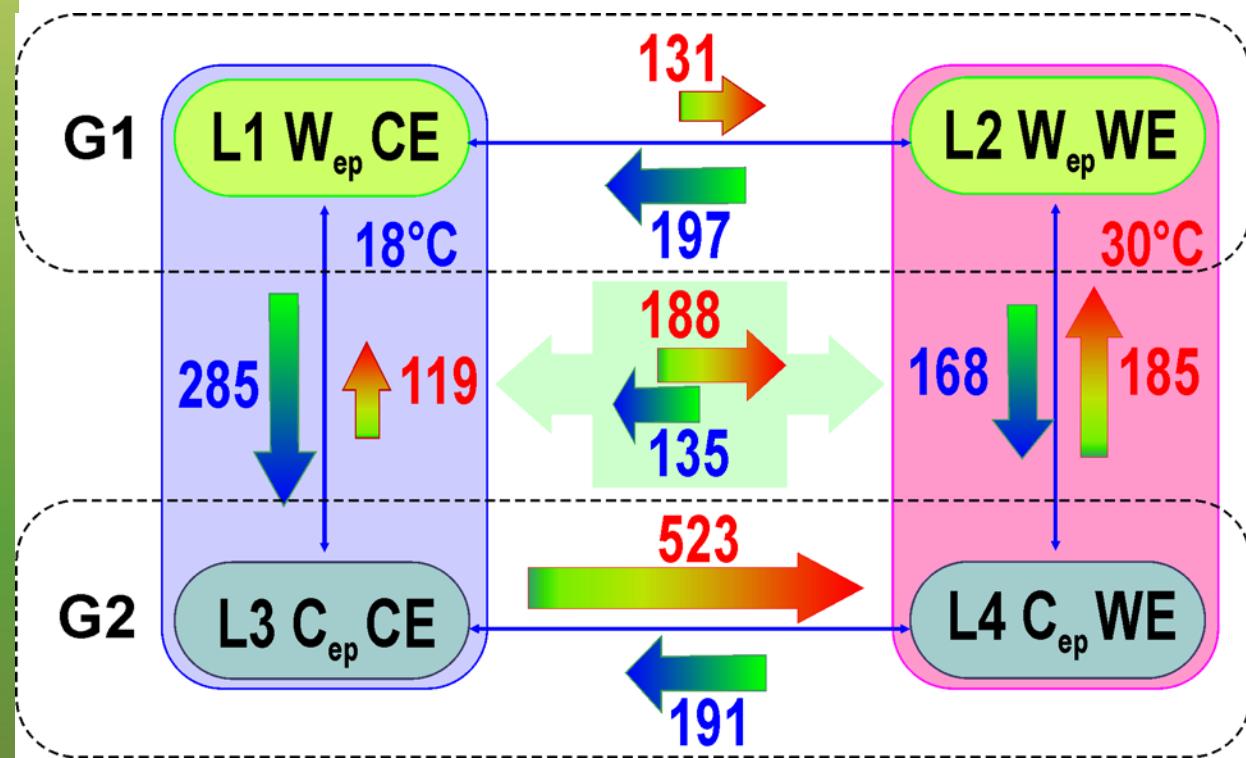


Gene ontology (GO) term assignments to Norway spruce unique transcripts ($p\text{-value} \leq 1e^{-10}$)



No functional differences between libraries, but large differences in GO terms between differentially expressed genes

Transcriptome differences (> 3 fold change)

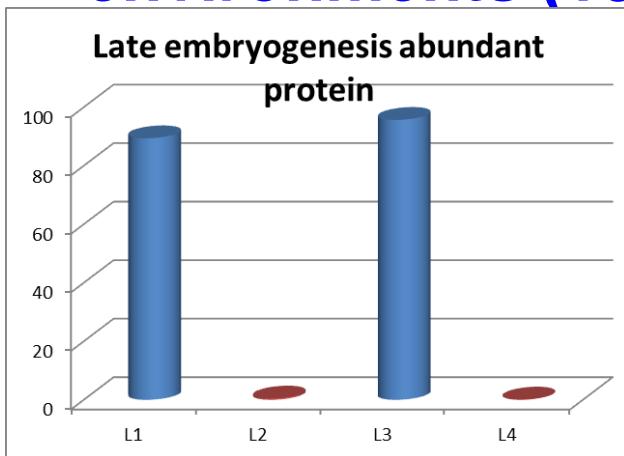


L1-L3 and L2-L4 shows transcription differences between warm and cold originated clones with the same treatment, i.e. reveal establishment of epigenetic memory marks (>323 DE genes)

L1-L2 and L3-L4 shows transcription differences between epitypes and related to family genotypic differences and putatively related epigenetic state transferred from previous “generation”, i.e. epigenetic memory

Gene expression differs between epitypes and treatments

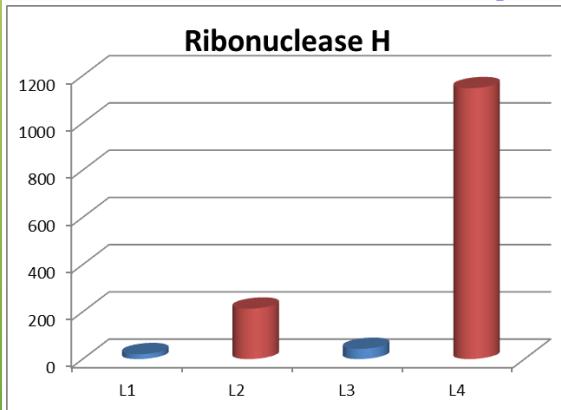
Selected most abundant transcripts in cold environments (18°C)



Tentative annotation	Average FC
Late embryogenesis abundant protein; n=1; <i>Picea glauca</i>	561,0
Trypsin inhibitor; n=1; <i>Bauhinia variegata</i>	158,6
F10A5.18 similar to glyoxal oxidase ; n=1; <i>Arabidopsis thaliana</i>	134,3
Peroxidase; n=1; <i>Pinus pinaster</i>	97,0
Trypsin inhibitor; n=1; <i>Bauhinia variegata</i>	93,0
PR-4 type protein; n=1; <i>Vitis vinifera</i>	84,5
Late embryogenesis abundant protein; n=1; <i>Picea glauca</i>	82,4
Pollen allergen CJP38; n=1; <i>Cryptomeria japonica</i>	55,2
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	44,2
DNA-cytosine methyltransferase; n=1; <i>Rhodopseudomonas palustris BisA53</i>	43,4
Putative uncharacterized protein; n=1; <i>Botryotinia fuckeliana B05.10</i>	30,0
Chalcone isomerase 4B; n=1; <i>Glycine max</i>	25,8
Predicted protein; n=1; <i>Physcomitrella patens subsp. patens</i>	25,0
Cell wall invertase; n=1; <i>Sorghum bicolor</i>	21,5
Putative uncharacterized protein; n=1; <i>Vitis vinifera</i>	19,2
Class IV chitinase Chia4-Pa2; n=2; <i>Picea abies</i>	18,6
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	17,9
Late embryogenesis abundant protein; n=1; <i>Picea glauca</i>	15,0
Nectarin IV; n=1; <i>Nicotiana langsdorffii</i> x <i>Nicotiana sanderae</i>	14,9
Leucoanthocyanidin dioxygenase; n=2; <i>Arabidopsis thaliana</i>	14,9
Envelope protein; n=1; Human immunodeficiency virus 1; Putative serine type endopeptidase inhibitor; n=1; <i>Zea mays</i>	14,1
Hydroxyproline-rich glycoprotein; n=1; <i>Glycine max</i>	14,0
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	13,4
tRNA-(M5C) methyltransferase; n=1; <i>Chlamydomonas reinhardtii</i>	11,1
Extensin; n=1; <i>Catharanthus roseus</i>	9,2

P-value= ~0

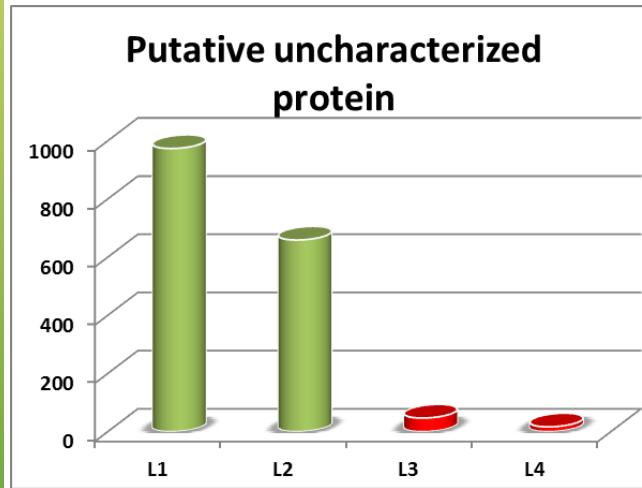
Selected most abundant transcripts in warm environments (30°C)



Tentative annotation	Average FC
Ribonuclease H; n=1; <i>Picea abies</i>	21,3
Cell wall-associated hydrolase; n=1; Roseobacter sp. AzwK-3b	30,4
Defensin; n=1; <i>Pinus sylvestris</i>	28,5
Retrotransposon protein; n=1; <i>Oryza sativa</i> Indica Group	21,3
WD-repeat protein-like; n=1; <i>Arabidopsis thaliana</i>	21,2
Metallothionein-like protein EMB30; n=2; <i>Picea abies</i>	19,3
Predicted protein; n=1; <i>Physcomitrella patens</i> subsp. <i>patens</i>	16,4
Cell wall-associated hydrolase; n=1; Roseobacter sp. AzwK-3b	12,6
NADH-ubiquinone oxidoreductase chain 1; n=7; core eudicotyledons	12,1
Heat shock protein; n=1; <i>Hevea brasiliensis</i>	11,1
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	10,9
un-annotated	9,4
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	9,2
Putative uncharacterized protein; n=1; <i>Oryza sativa</i> Japonica Group	9,1
un-annotated	9,1
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	8,9
Glycine-rich RNA-binding protein; n=1; <i>Picea glauca</i>	7,6
hypothetical protein MTR_5g051120; ATP synthase subunit beta [Medicago truncatula]	6,7
Gibberellin regulated protein; n=1; <i>Medicago truncatula</i>	6,4
Serine/threonine protein phosphatase; n=1; <i>Ostreococcus tauri</i>	6,2
Ribonuclease H; n=1; <i>Picea abies</i>	6,2
Cytochrome P450 like_TBP; n=1; <i>Nicotiana tabacum</i>	6,0
MYB92; n=1; <i>Malus x domestica</i>	6,0
Anthocyanidin synthase-like protein; n=1; <i>Arabidopsis thaliana</i>	
Putative uncharacterized protein orf111-b; n=1; <i>Zea mays</i>	5,7

P-value= ~0

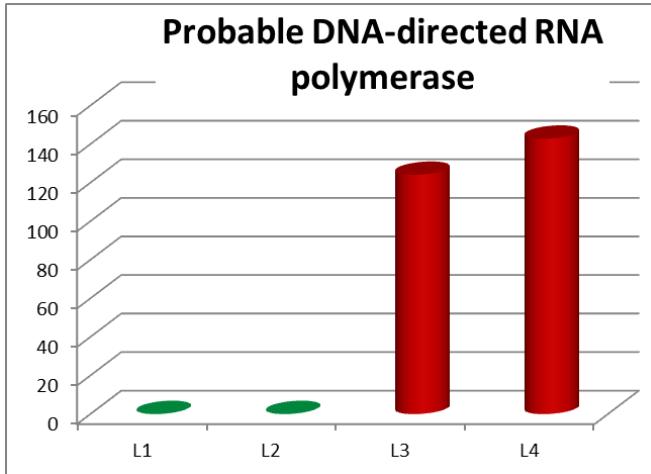
Selected most abundant transcripts in warm epitype



Tentative annotation	Average FC
Polyubiquitin - <i>Fragaria ananassa</i>	14238,9
Ubiquitin - <i>Arabidopsis thaliana</i>	3891,6
Putative uncharacterized protein	3301,1
Antimicrobial peptide 1; n=1; <i>Pinus sylvestris</i>	5158,5
Predicted protein; n=5; <i>Physcomitrella patens</i> subsp. <i>patens</i>	3742,2
Arabinogalactan-like protein; n=1; <i>Pinus taeda</i>	3545,9
un-annotated	2899,5
60S ribosomal protein L41 - <i>Arabidopsis thaliana</i>	2171,6
60S ribosomal protein L27a-3; n=2; <i>Arabidopsis thaliana</i>	1999,9
un-annotated	1454,6
60S ribosomal protein L41; n=10; Magnoliophyta	1423,1
Septin-14; n=1; <i>Mus musculus</i>	1222,5
un-annotated	1095,4
Predicted protein; n=1; <i>Physcomitrella patens</i> subsp. <i>patens</i>	1039,6
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	1016,1
Putative uncharacterized protein; n=1; <i>Picea abies</i>	1002,4
Glyceraldehyde-3-phosphate dehydrogenase precursor; n=1; <i>Pinus sylvestris</i>	529,1
Cytochrome c oxidase, cbb3-type, subunit II; n=1; <i>Loktanella vestfoldensis SKA53</i>	472,1
un-annotated	435,6
Alpha-L-arabinofuranosidase-like protein precursor; n=1; <i>Victivallis vadensis</i> ATCC BAA-548	221,7
Methylenetetrahydrofolate reductase; n=1; <i>Vitis vinifera</i>	184,3
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	176,1
Antimicrobial peptide 1; n=1; <i>Pinus sylvestris</i>	161,8
Alpha-dioxygenase; n=1; <i>Pisum sativum</i>	160,5
60S ribosomal protein L39; n=6; Poaceae	144,4

P-value= ~0

Selected most abundant transcripts in cold epitype



Tentative annotation	Average FC
Expressed protein; n=1; <i>Arabidopsis thaliana</i>	3418,8
Arabinogalactan-like protein - <i>Pinus taeda</i> (Loblolly pine), partial (26%)	2968,5
Probable DNA-directed RNA polymerase subunit delta; n=1; <i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363	2669,8
Putative uncharacterized protein; n=1; <i>Ustilago maydis</i>	2418,9
Weak similarity to tRNA modification GTPase TrmE	1641,5
un-annotated	1544,3
KH domain-containing protein-like; n=2; <i>Oryza sativa</i>	1480,4
Predicted protein; n=1; <i>Physcomitrella patens</i> subsp. <i>Patens</i>	1408,2
Putative pollen specific protein; n=1; <i>Oryza sativa</i>	1398,6
Cytochrome c oxidase polypeptide VI, mitochondrial precursor; n=1; <i>Neurospora crassa</i>	1352,5
Probable glutathione S-transferase; n=1; <i>Picea mariana</i>	1286,2
Translation elongation factor-1 alpha; n=3; <i>Pinaceae</i>	1212,2
un-annotated	1123,8
Non-cell-autonomous heat shock cognate protein 70; n=1; <i>Cucurbita maxima</i>	1115,5
Nodule protein Dg93-like protein; n=1; <i>Potamogeton distinctus</i>	1094,2
un-annotated	1085,7
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	1071,8
Putative uncharacterized protein; n=1; <i>Picea sitchensis</i>	902,4
un-annotated	364,4
Pathogenesis-related protein 4b; n=1; <i>Capsicum chinense</i>	331,2
Fasciclin-like arabinogalactan protein 11; n=1; <i>Gossypium hirsutum</i>	319,3
Putative 60S ribosomal protein L1; n=1; <i>Trifolium pratense</i>	241,7
Similar to gibberellin-stimulated transcript 1 like protein; n=1; <i>Arabidopsis thaliana</i>	239,7
Extensin class 1 protein precursor; n=2; <i>Vigna unguiculata</i>	skog 224,5 landskap
Arabinogalactan-like protein; n=1; <i>Pinus taeda</i>	206,9

P-value= ~0

Conclusions

High-throughput sequencing confirmed striking differences in transcriptomes of the same epitypes growing in different temperatures and between epitypes

We can judge that warm and cold epitypes react various ways and activate different set of genes in CE and WE but also react similar ways leading to novel establishing of epigenetic memory in embryos developing in CE and WE.

Future work

Work in Progress

Analyze transcriptomes during other stages of embryo development *in vitro*

Analysis of chromatin state of candidate genes (in the frames of ProCoGen)

Analysis of methylation state (general and gene specific)

Norwegian Forest and Landscape Institute

Igor Yakovlev

Carl Gunnar Fossdal

Tore Skrøppa

Harald Kvaalen

Anne Nilsen



Norwegian University of life Sciences

YeonKyeong Lee

Jorunn Olsen

Øystein Johnsen



GenXPro GmbH

Björn Rotter



The Research Council
of Norway