



СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
SIBERIAN FEDERAL UNIVERSITY



ЦЕНТР
ГЕНОМНЫХ
ИССЛЕДОВАНИЙ
научно-образовательный центр СФУ

HOMEBOX GENES: INVESTIGATING THE DEVELOPMENT OF *PINUS SYLVESTRIS* (SCOTS PINE)

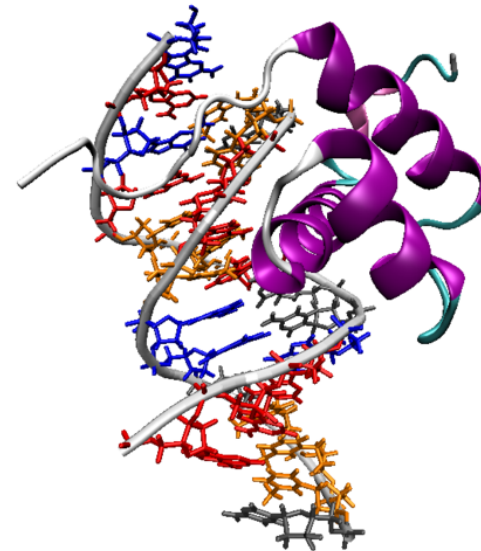
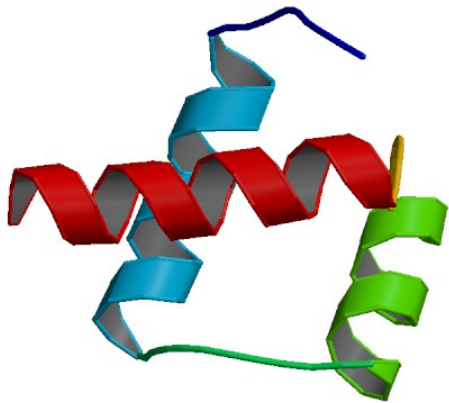
Tatiana Guseva, Vladislav Biriukov, Michael Sadovsky

- Comprehensive gene expression profiling of homeobox gene family members in *P. sylvestris* **provides opportunity for functional analysis to reveal their exact role in plant growth and development**
- With sufficient **understanding** of homeobox genes **functions**, it is easier to investigate another phenomenae like plant **response to abiotic and biotic stress**.

Introduction. Homeobox



- The **homeobox** encodes the **homeodomain** currently found in many **developmental control genes**.
- **Homeobox genes** encode **transcriptional factors** and most of them play important role in **development** of multicellular organisms.
- Plant homeodomain proteins can be split into **14 distinct classes**



Images:

Clarke, Neil D., et al. "Structural studies of the engrailed homeodomain." *Protein Science* 3.10 (1994): 1779-1787.

Billeter, Martin, et al. "Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex." *Journal of molecular biology* 234.4 (1993): 1084-1097.

- The **aim** of the work is to examine the development of Scots pine (*Pinus sylvestris* L.) at the homeobox-containing genes level using the differential gene expression profile.

To do that we have done:

- Screening of genomic resources to identify RNA-seq source from different tissues of *P. sylvestris*;
- Pre-processing of the sequences: assessment of quality of reads;
- *De novo* transcriptome assembly;
- Selection of homeobox-containing transcripts;
- Differential expression analysis and functional annotation.

Material and methods



- *P. sylvestris* sequencing data are deposited in the NCBI BioProject database provided by the Norwegian Institute of Bioeconomy Research
- The material comprises **five tissues** (needle, phloem, vegetative bud, embryo and megagametophyte)



Pinus sylvestris



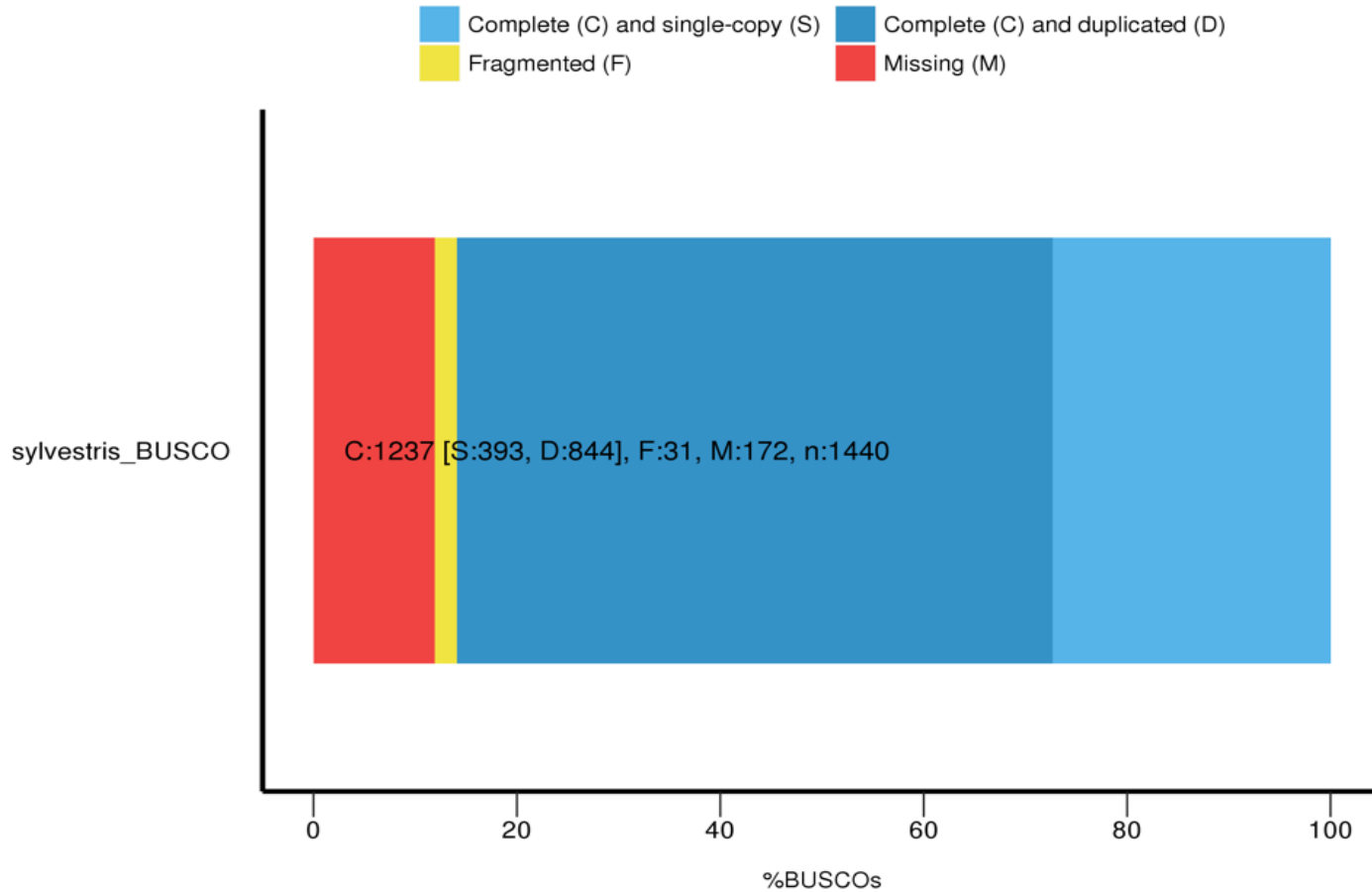
Methods:

- **HMMER** software (version 3.2.1) for the homeobox domain identification
- *De novo* transcriptome assembly with **Trinity** software (version 2.8.4)
- Differential expression analysis and functional annotation (FA) (**RSEM**, **Bowtie** programs (version 1.2.3), **EdgeR** package (**R** version 3.5.0, Bioconductor version 3.8), **UniProt** database for FA)



Total trinity 'genes'	488 116
Total trinity transcripts	775 502
GC%	40.19
Stats based on ALL transcript contigs	
Contig N50 (bp)	1 273
Median contig length (bp)	360
Average contig (bp)	713.6
Total assembled bases	553 398 248
Stats based on the LONGEST ISOFORM per 'gene' only	
Contig N50 (bp)	596
Median contig length (bp)	313
Average contig (bp)	513.56
Total assembled bases	250 675 880

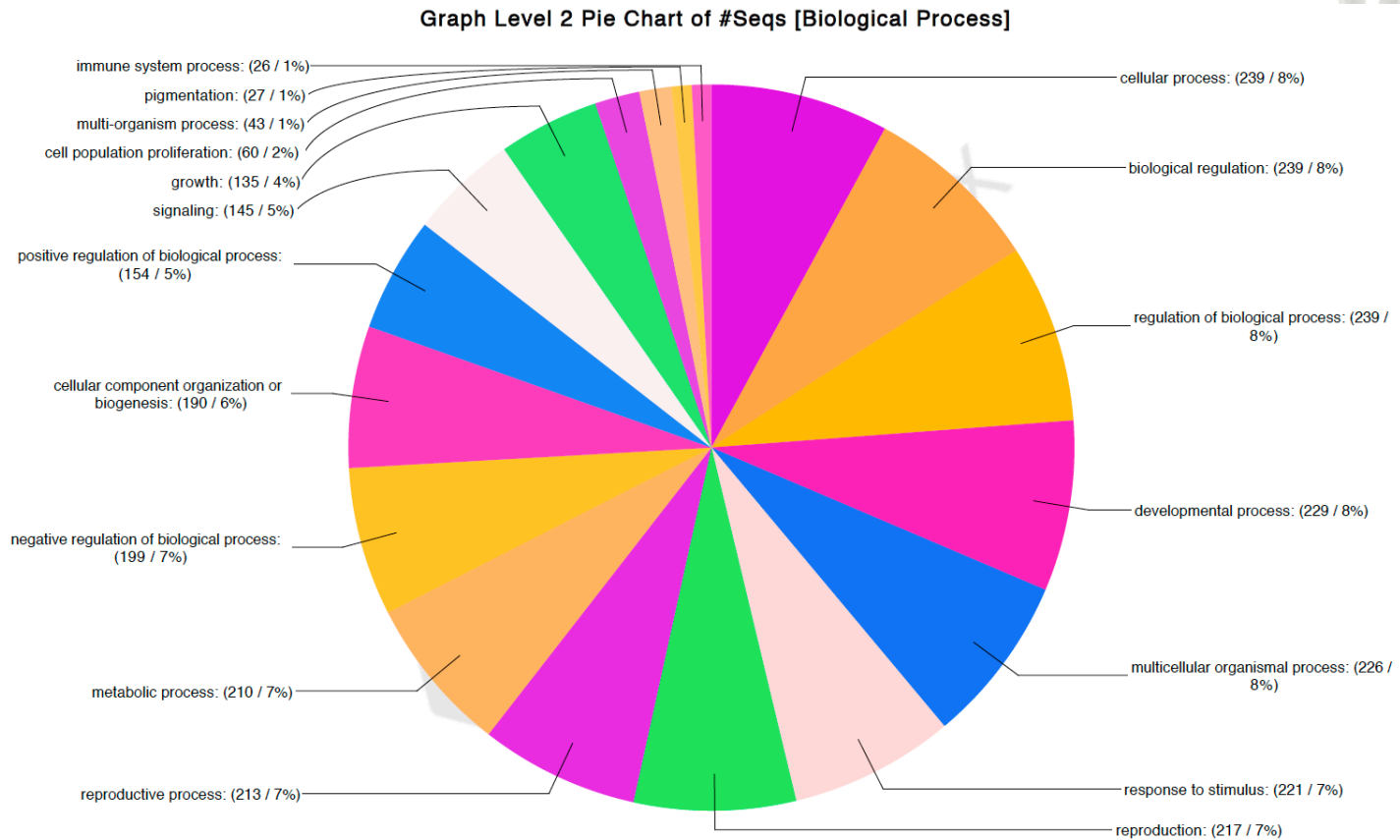
BUSCO Assessment Results



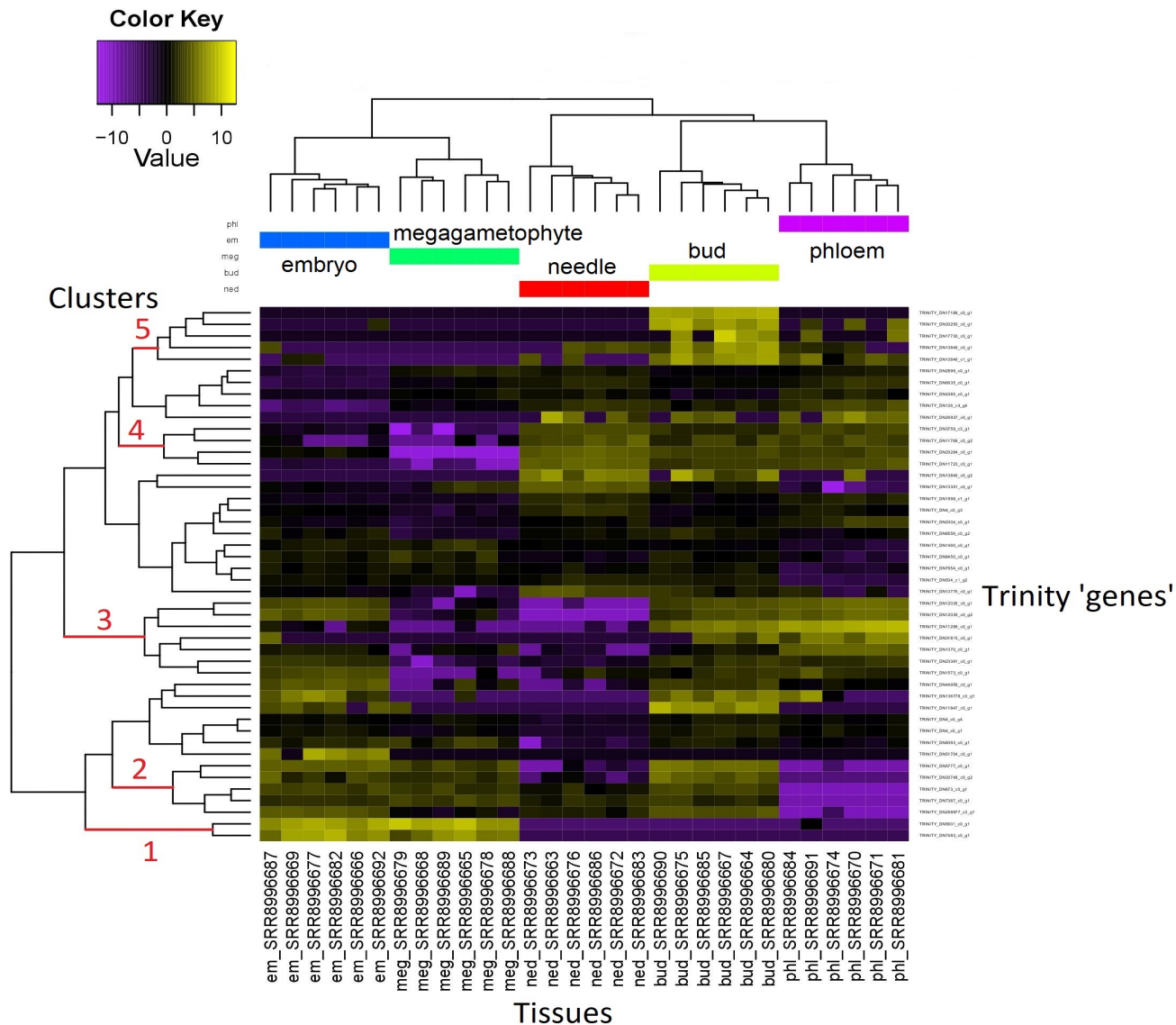


Classification of homeobox gene family members in *P. sylvestris*

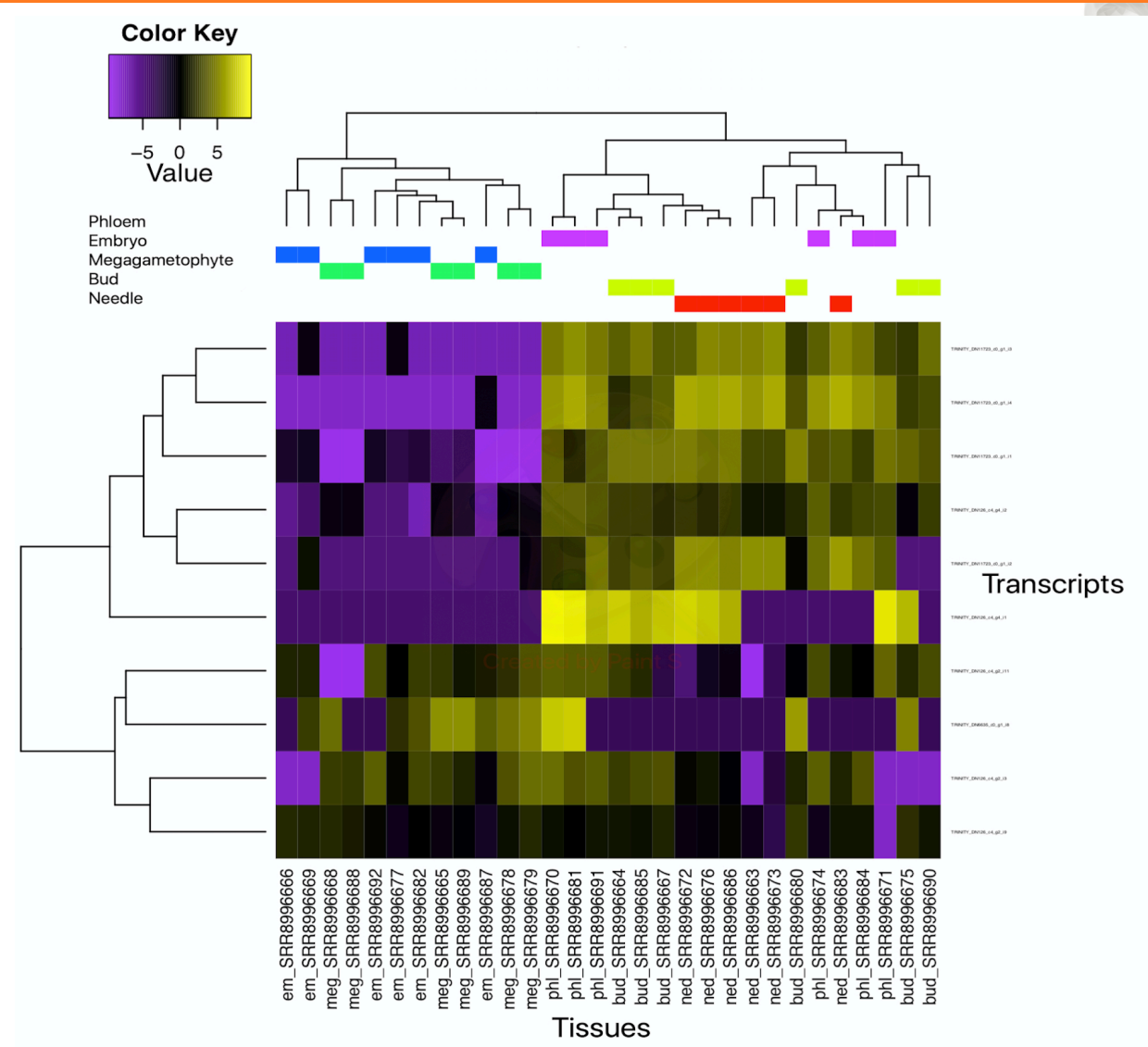
<i>P.Sylvestris</i>		<i>P.Sylvestris</i>	
Class	Number of transcripts	Class	Number of transcripts
HD-ZIP I	67	DDT	11
HD-ZIP II	29	PHD	4
HD-ZIP III	15	PINTOX	0
HD-ZIP IV	26	LD	4
PLINC	1	NDX	0
WOX	29	SAWADEE	11
BEL	21	Unclassified	14
KNOX	11	Total	243



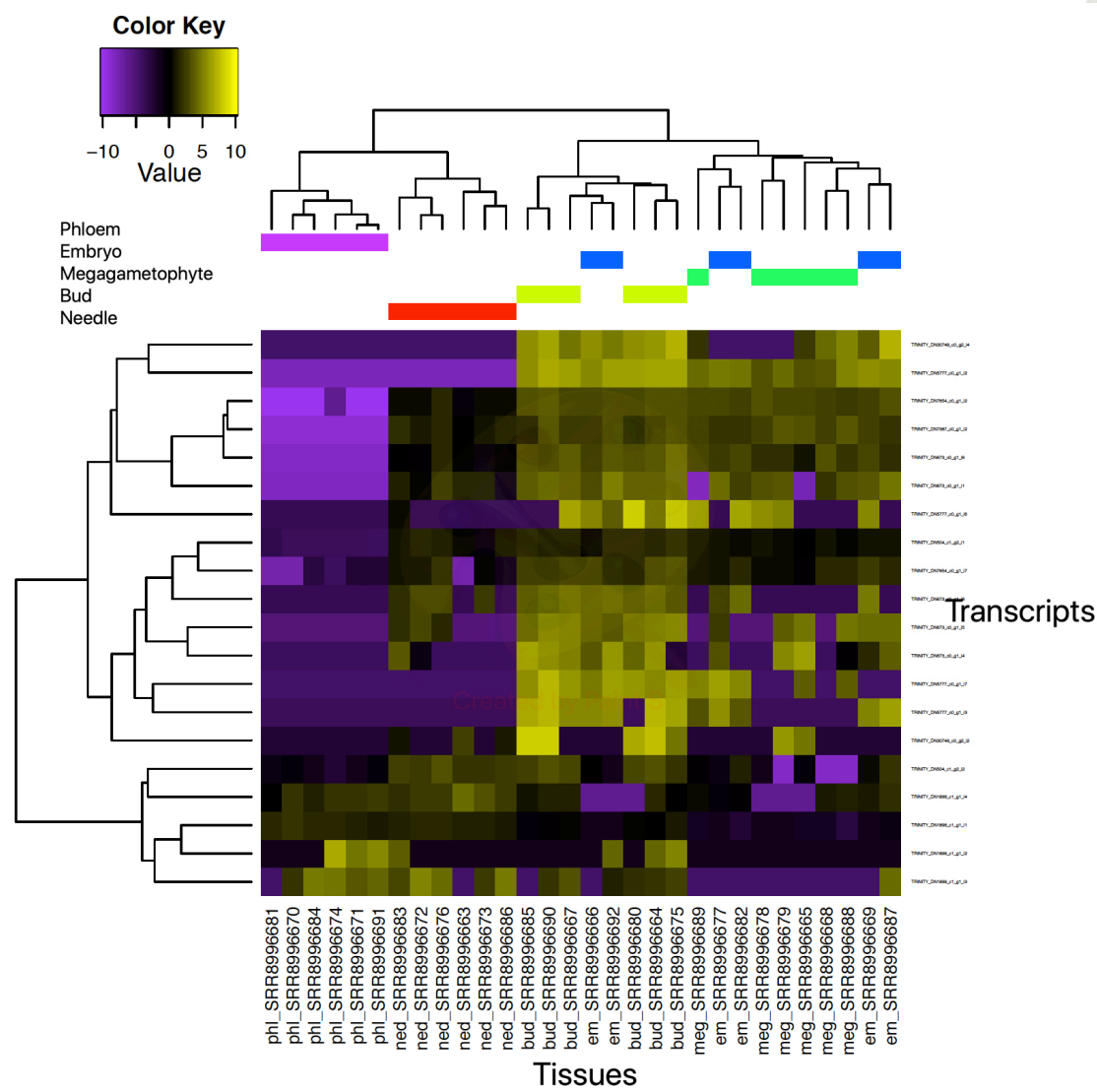
Distribution of biological functions of homeodomains



Heatmap of gene expression, five tissues of *Pinus sylvestris*



Heatmap of gene expression of BEL class homeobox genes, five tissues of *Pinus sylvestris*



Heatmap of gene expression of HD-ZIP IV class homeobox genes, five tissues of *Pinus sylvestris*

- *De novo* transcriptome assembly shows the higher numbers of BUSCO percentage of transcriptome completeness in comparison to previously reported results.
- 243 transcripts containing homeobox genes were selected and annotated yielding 14 classes of homeobox families (*2 of them were not identified*).
- Five clusters of DE transcripts were analyzed.
- The members of BEL class were found to be up-regulated in needles, phloem and bud and down-regulated at megagametophyte and embryo, thereby suggesting their role in development of tissues of the mature tree.
- Expression patterns of HD-ZIP IV genes showed considerable down-regulation in phloem and needles.

Thank you for attention!



- **Cluster 1.** Two transcripts encoding proteins involved in early and late embryo development were found, which were upregulated at embryo and megagametophyte stages and down-regulated at other stages.
- **Cluster 2.** Cluster comprises transcripts encoding transcriptional factors (TF) involved in seed and leaf development as well as in cell differentiation, down-regulated in phloem tissue.
- **Cluster 3.** Transcripts annotated as proteins responsible for multicellular organism development, phloem histogenesis, auxin-mediated morphogenesis, plant organ development, regulation of leaf morphology and involved in meristem formation were observed. These transcripts are down-regulated in needle and megagametophyte tissues.
- **Cluster 4.** Cellular response to cytokine stimulus, detection of cytokinin stimulus, response to light stimulus and BEL1-like homeodomain 4(BLH4) which is a transcription factor that establishes leaf shape by repressing growth in specific subdomains of the leaf and negatively regulates knox homeobox gene KNAT1/BP expression. Up-regulated at needle, bud and phloem.
- **Cluster 5.** Transcripts in cluster 5 included proteins involved in the cotyledon morphogenesis, cell differentiation, leaf morphogenesis. Transcripts of this cluster are significantly up-regulated in bud and down-regulated in other tissues.