



SCIENCE AND
EDUCATION **FOR**
SUSTAINABLE
LIFE

The dark matter of the spruce genome

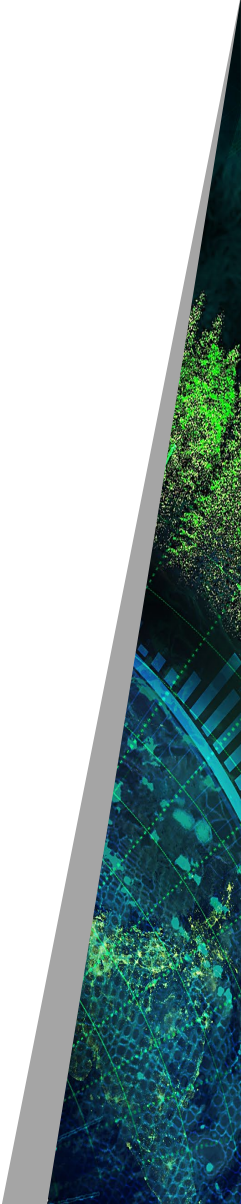


Peter Kindgren

Umeå Plant Science Centre

Swedish University of Agricultural Sciences

B4E meeting, 22nd October 2023



Back to the basics

I'll be talking about spruce – applicable to all organisms

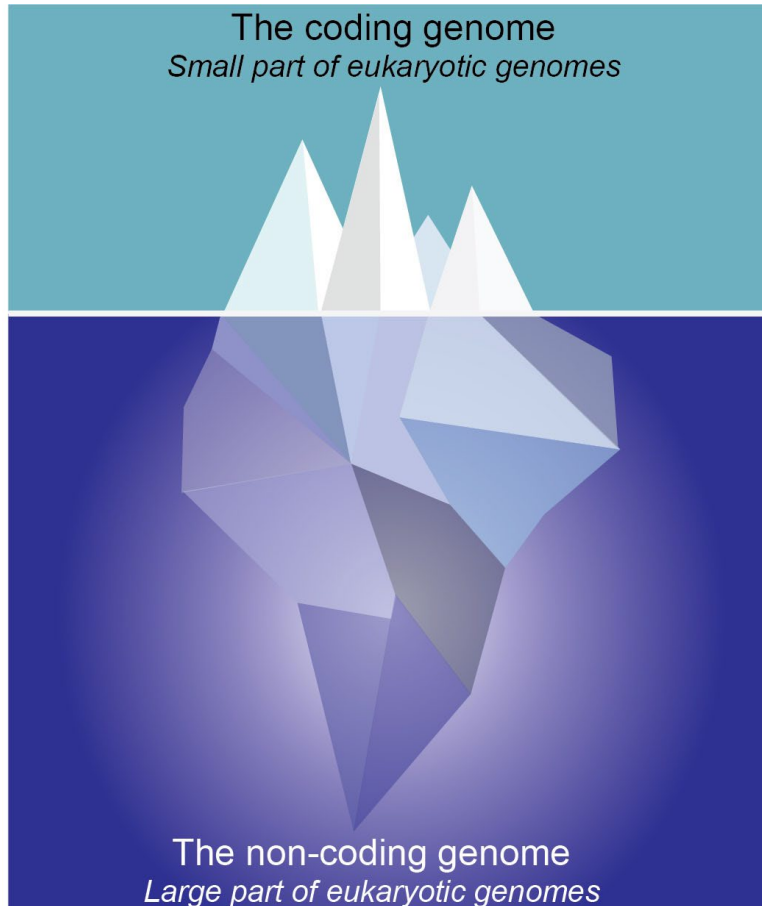
To maximize the output, we need to understand the basics

My interest: How do plants use their DNA?

What parts of the DNA is transcribed into RNA?

What is the function of non-coding RNA?

Transcription occurs everywhere in the genome



Stable, Expression level high

Unstable, Expression level low
(The dark matter)

Combining DNA sequence with RNA biology approaches

How do trees use their DNA blueprint to achieve plasticity?

We have massive amount of DNA sequence

How much is used as template for RNA?

- Transcription Start Site sequencing (TSS-seq)



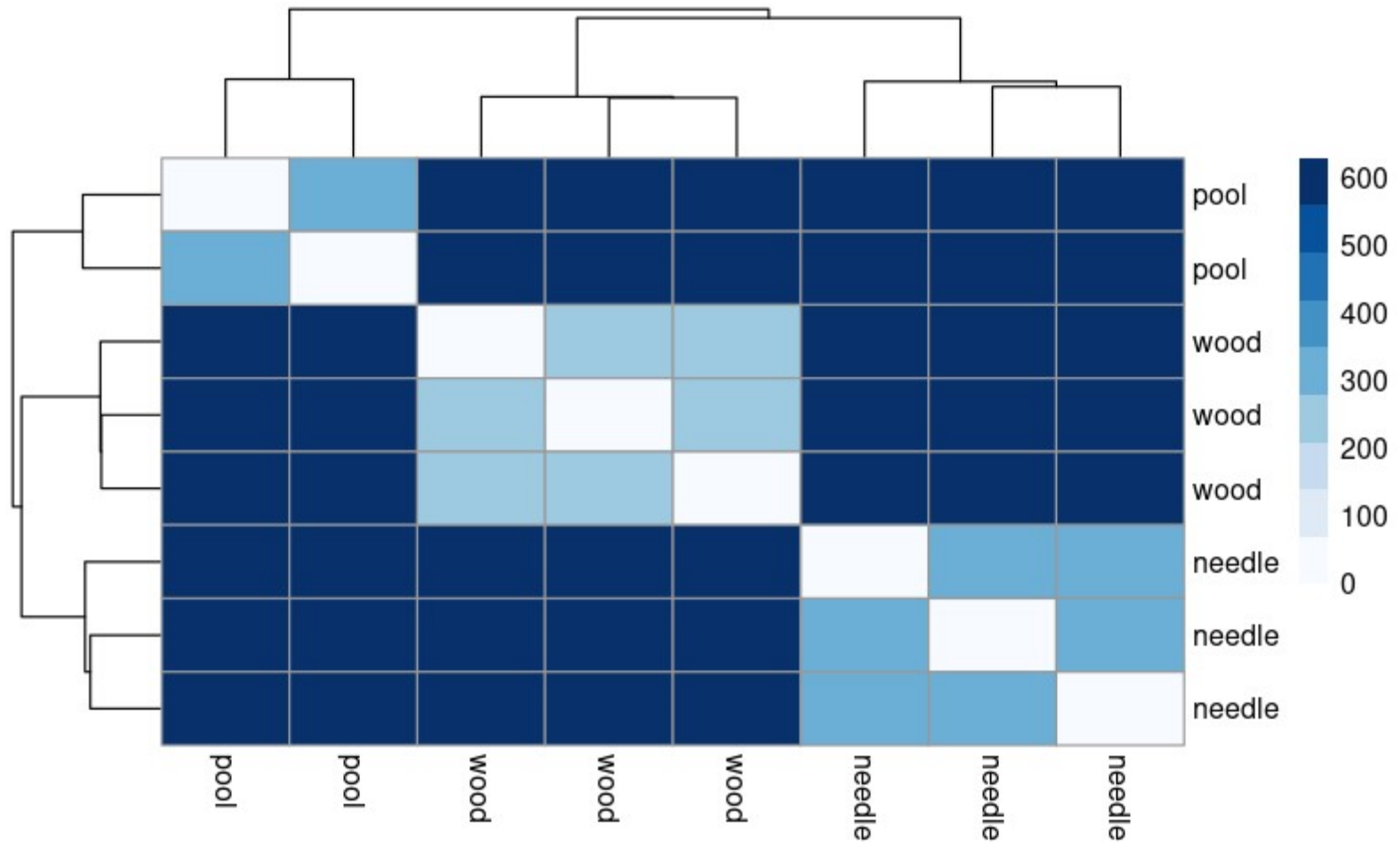
Wikipedia

TSS usage in different tissues of spruce samples

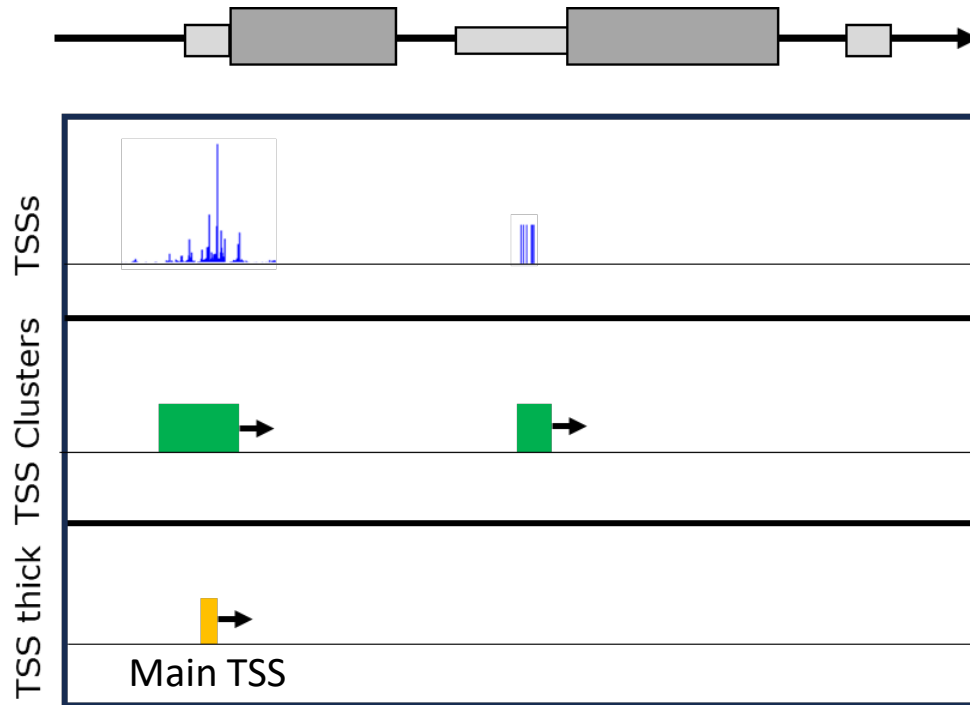
Spruce

- Pooled samples (PEM, embryo, xylem and pollen)
- Needle
- Wood (cambium and xylem)

TSS usage in different tissues of spruce samples



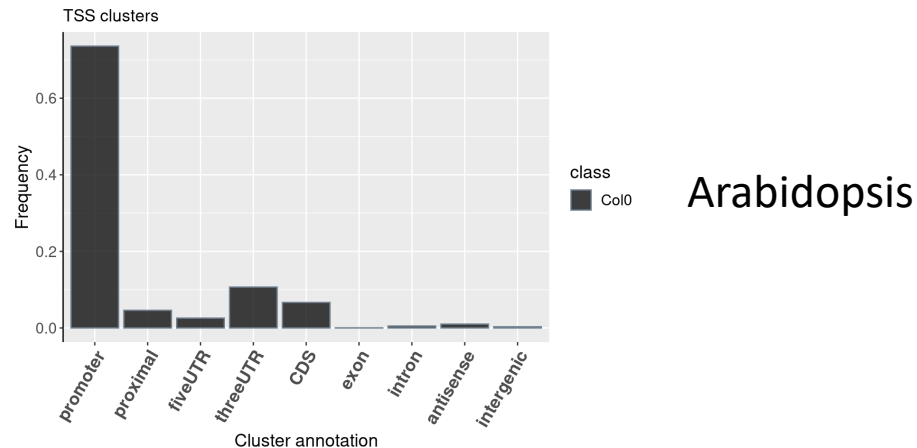
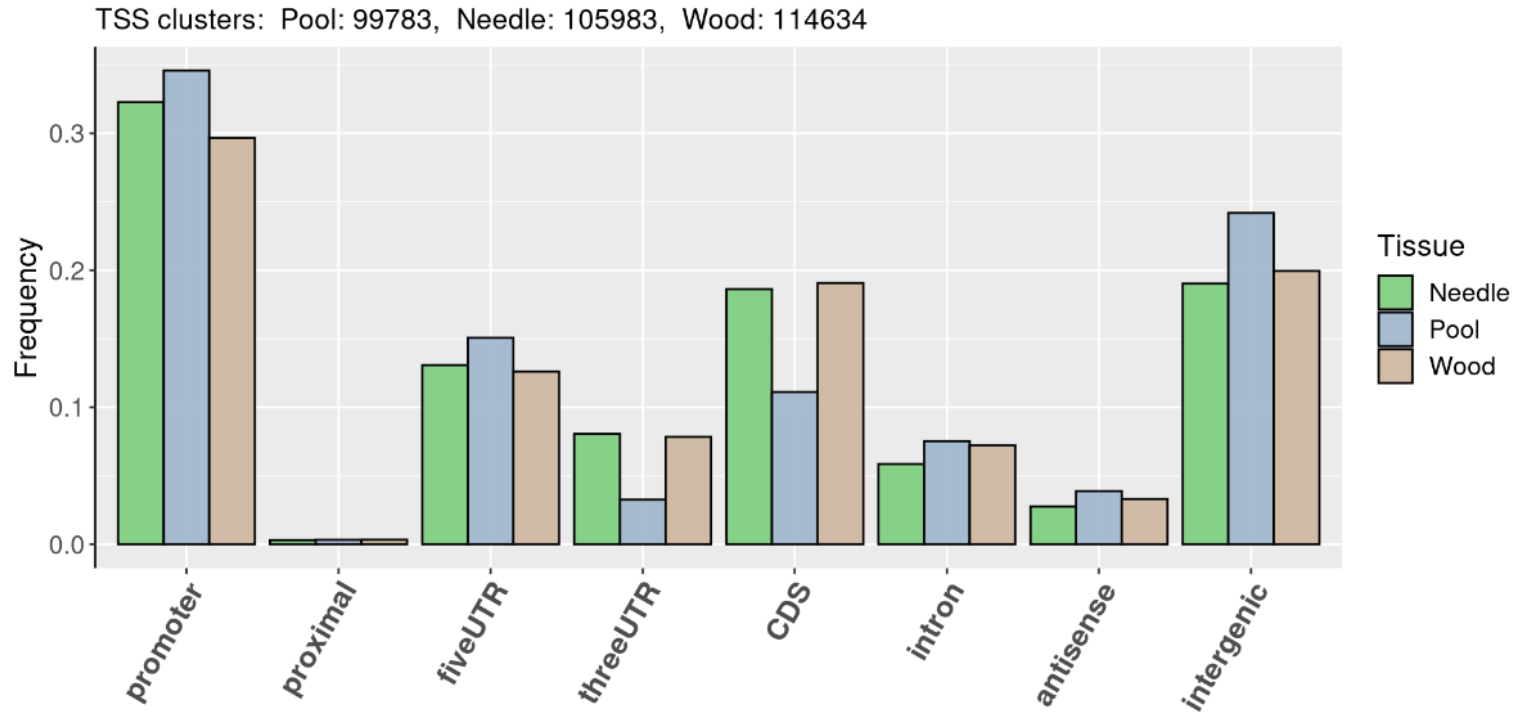
TSS usage in different tissues of spruce samples



12.917 million total TSS reads correspond to
29,183 annotated genes

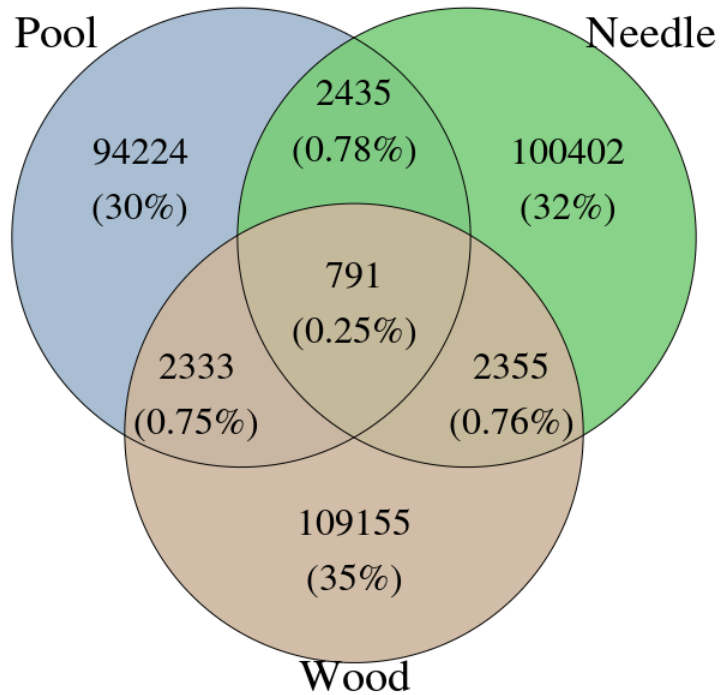
278,129 TSS clusters (TSSs within 50 bp) (TPM ≥ 1)
in at least two replicates

TSS clusters within spruce genome annotation categories

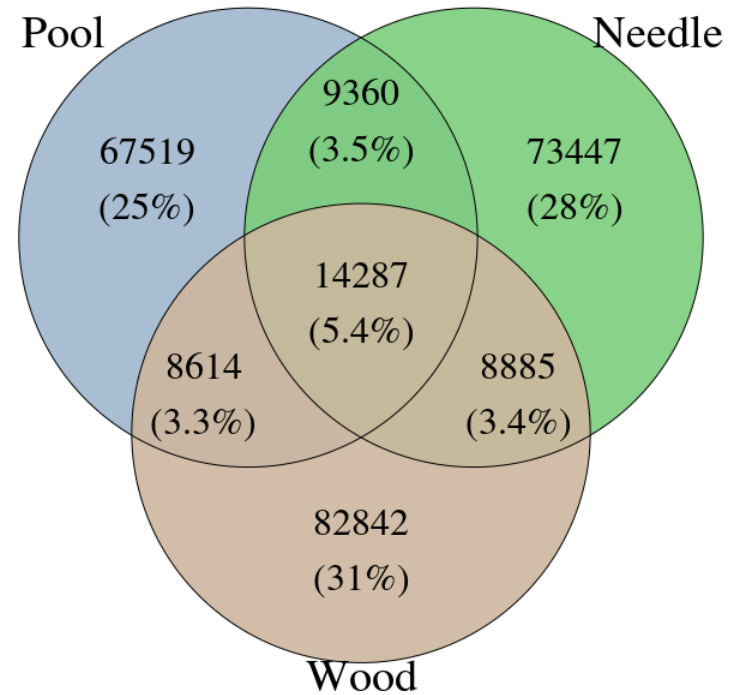


Cluster level analysis of TSS-seq in spruce

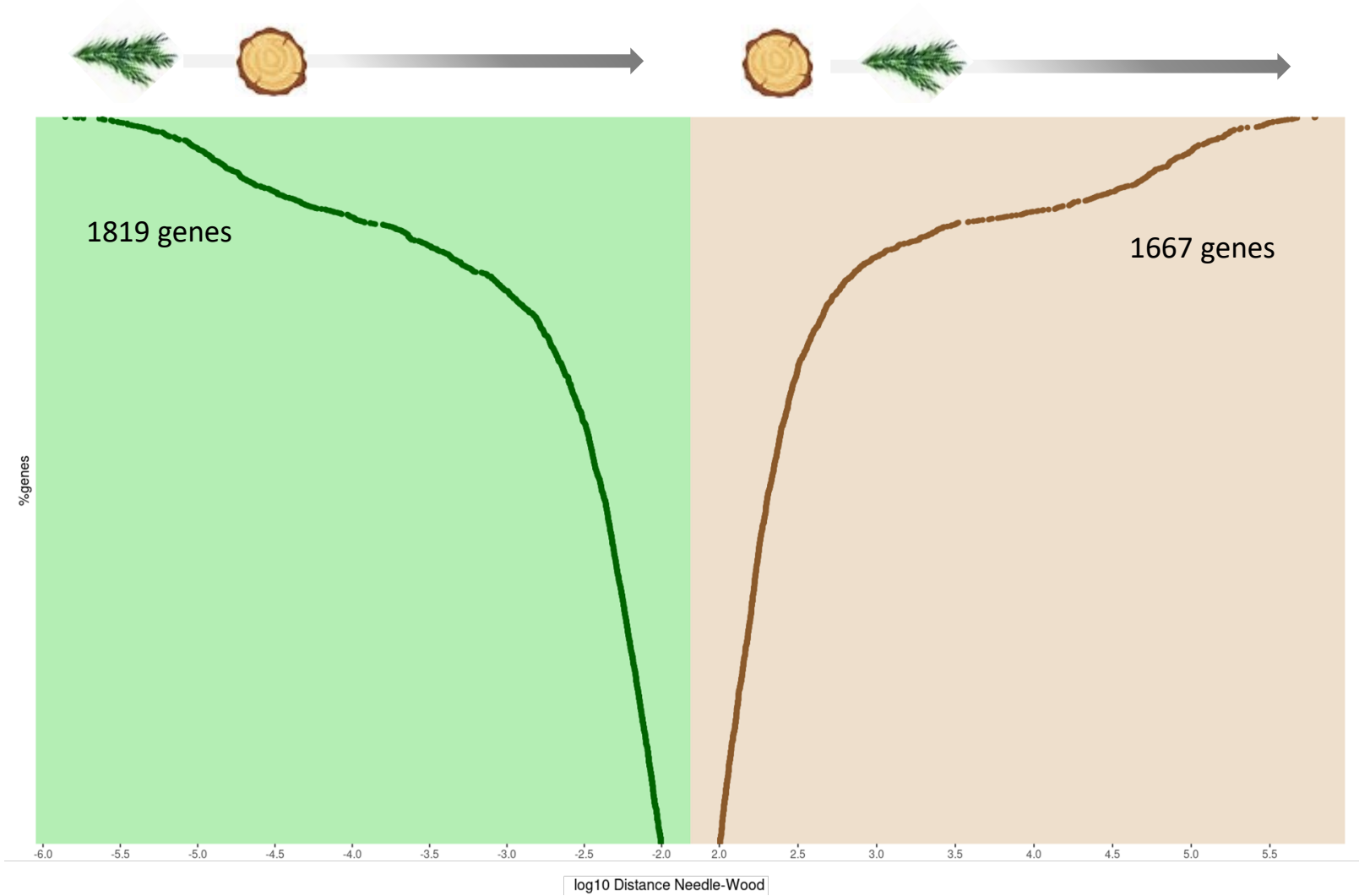
TSS clusters



Main TSS



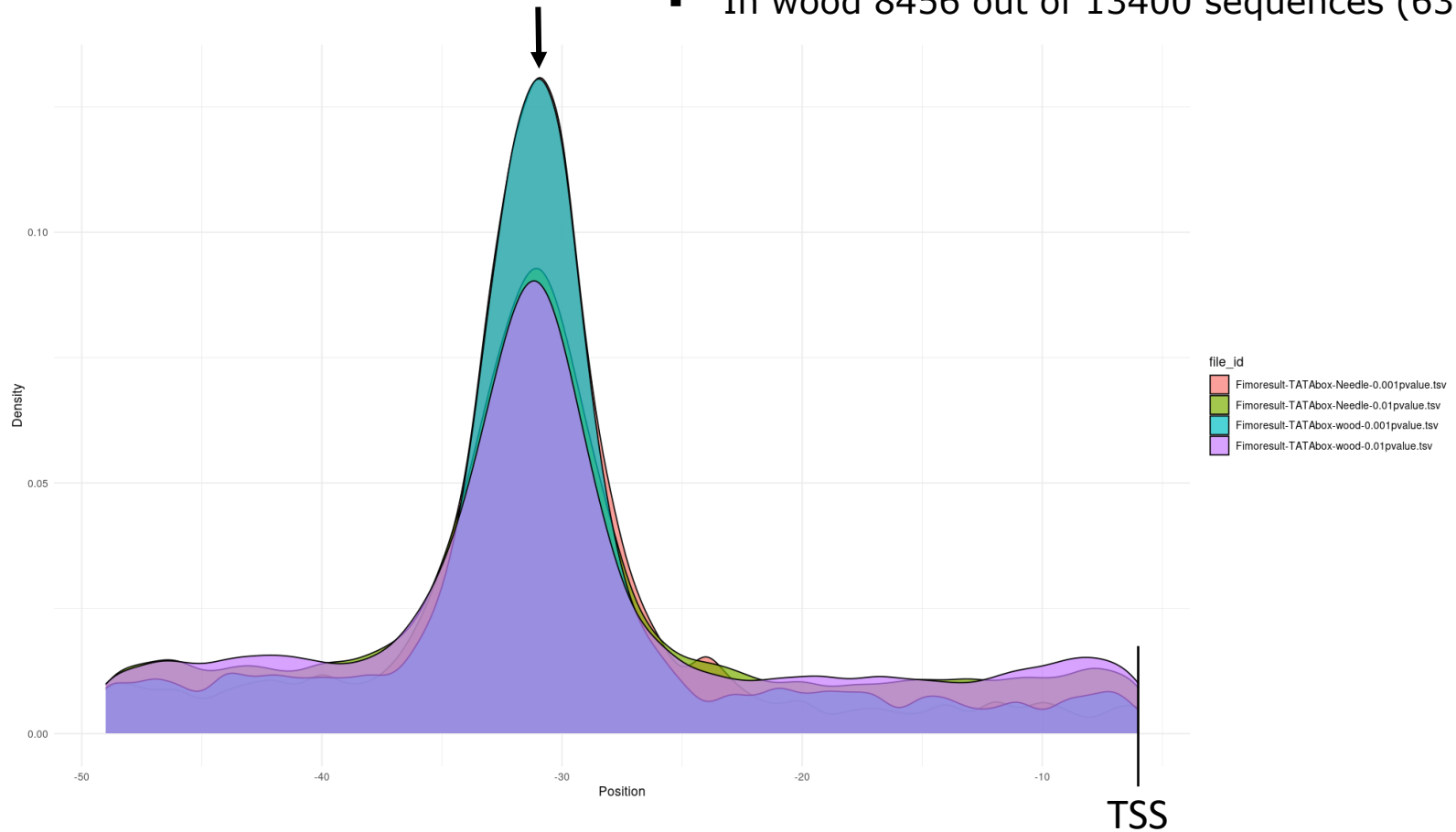
Distance analysis between TSSs in needle and wood



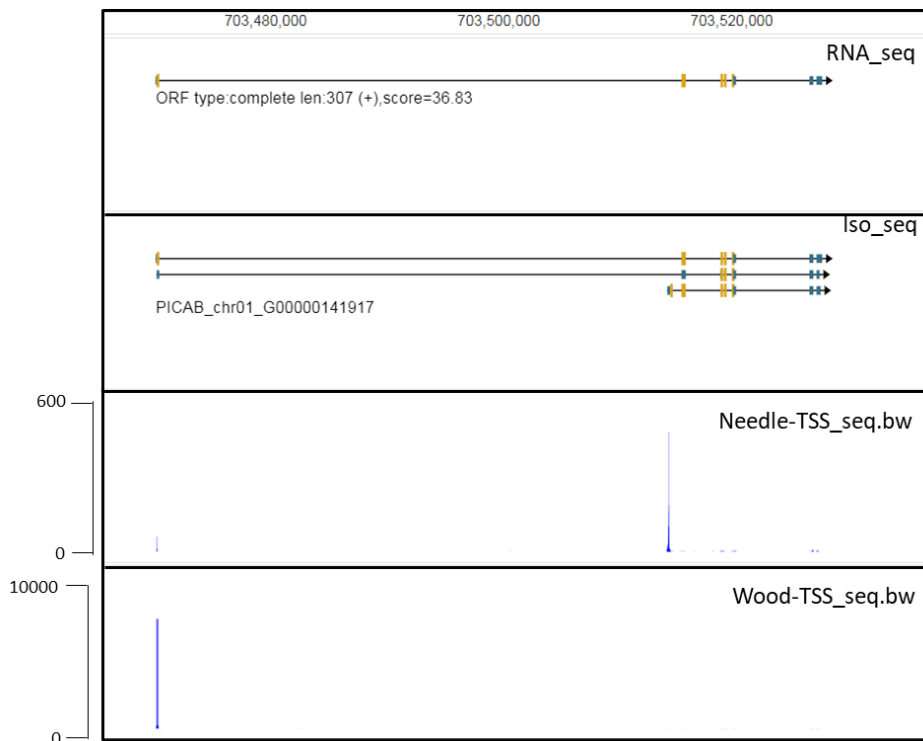
Different tissues use different promoters

p -value < 0.01

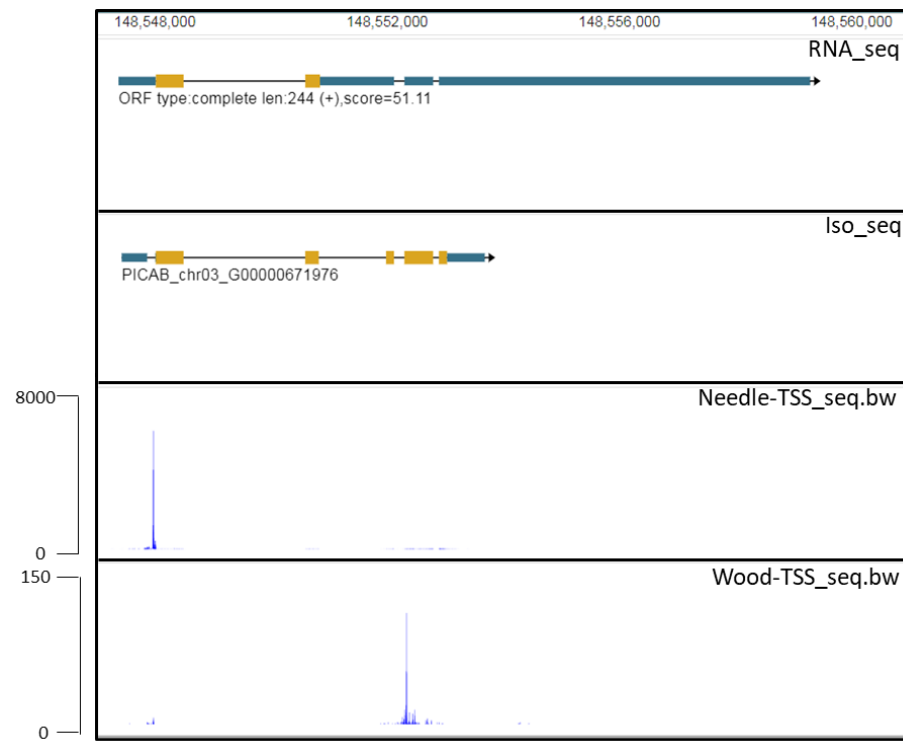
- In needle 8938 out of 13650 sequences (65%)
- In wood 8456 out of 13400 sequences (63%)



Alternative transcription start site usage



PICAB_chr01_G00000141917, PA_chr01:703,470,566..703,527,750 (+)



PICAB_chr03_G00000671976, PA_chr03:148,542,445..148,562,615 (+)

The next steps

Does spruce have tissue specific protein isoforms?

- Ribo-seq, proteomics

What dictates the use of specific promoters?

- Promoter analysis, DNA accessibility

Can we use this to create more biomass/better trees?

Acknowledgements

Vikash Kumar
Aman Zare
Theerarat Fai Kochakarn

