Population genomics of *Picea rubens*

Thomas O'Leary

February 25, 2020

Background

Picea rubens Ecology

In the context of a rapidly changing climate many organisms will be forced to either adapt, migrate, or go extinct. In the case of *Picea rubens*, a red spruce that thrives in cool and moist climates, this may lead to futher range contraction or eventual extinction if the species is unable to adapt in time. As glaciers began to melt approximately 20,000 years ago, *Picea rubens* was forced to retreat to isolated populations along the mountain tops of the Appalacian range. As a result, these populations along the mid-Atlantic United States are highly fragmented and completely isolated from the northen core of the range in Northern New England and Canada.

As the climate continues to warm, these isolated populations may pop out of existence due to the increased environmental stress. However, because evolution can only act on standing genetic variation these isolated populations may represent an important genetic resource for the species at large, and may help inform conservation efforts. The ultimate aims of this study are to (i) describe the population structure and genetic diversiy of *Picea rubens* along its current range, (ii) identify loci that show signs of positive selection and (iii) map the genetic basis of these adaptive phenotypes. In this write up, I will only begin to address the first aim.

Sample collection, library preparation, and sequencing

Whole genomic DNA was extracted from needle tissue that was collected from a total of 340 mother trees in 65 populations, of which 110 trees were from 23 edge populations. Probes were designed (80,000 120bp) for exome capture based on the multiple developmental and tissue type transcriptomes from the related white spruce *Picea glauca* (Rigault et al. 2011; Yeaman et al. 2014). Approximately 95% of the probes were designed within exomic regions with the remaining 5% in intergenic regions, covering a total of 38,570 unigenes. The probes were blasted against the *P. glauca* reference genome to ensure at least 90bp of 85% identity. 250ng to 1 μg of genomic DNA was mechanically sheared to an average length of 400 bp, exome fragments were enriched using designed probes and following barcode adaptation the libraries were pooled and paired-end 150 bp sequenced on an Illumina HiSeq X.

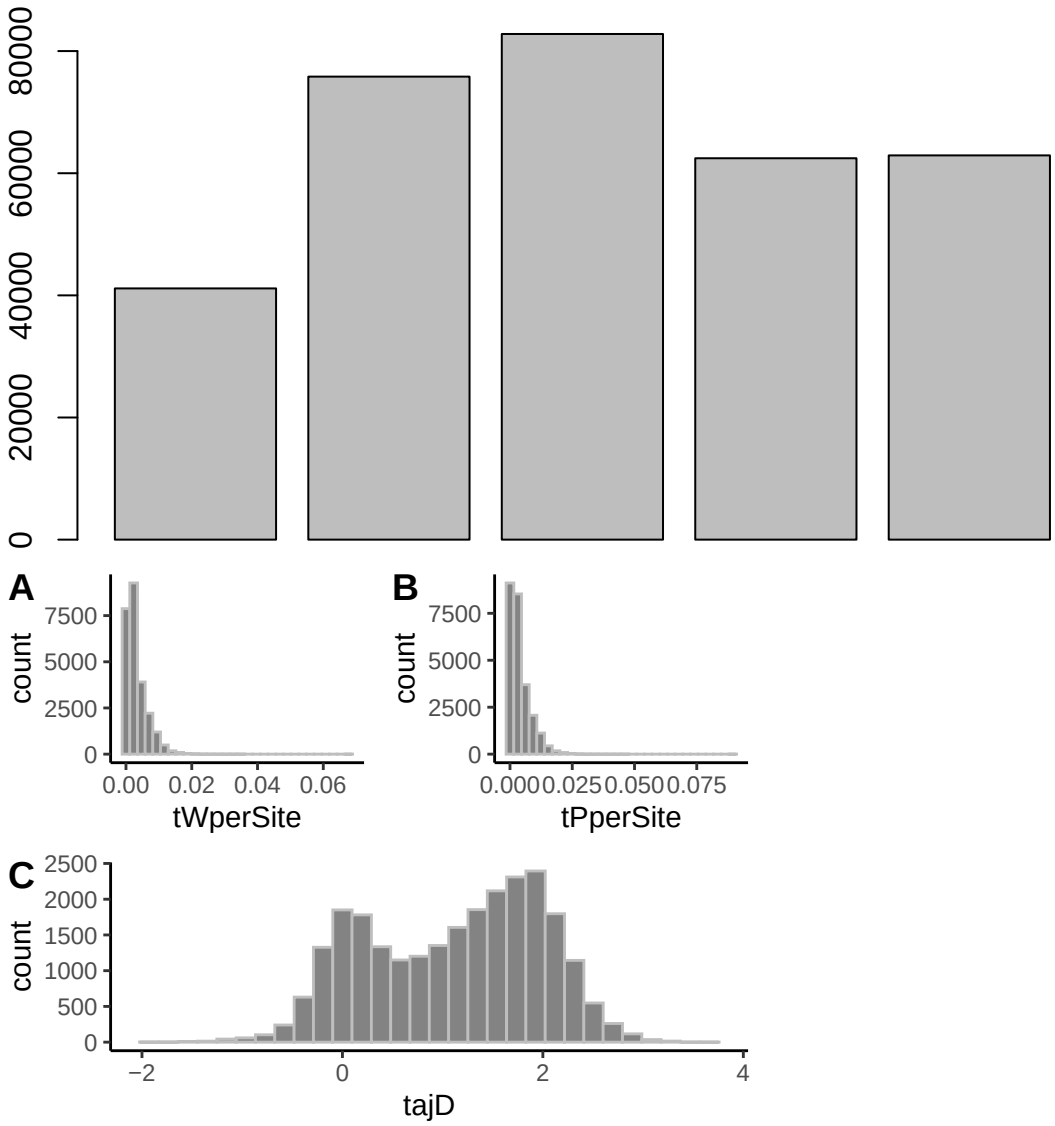
Bioinformatics Pipeline

The reads were mapped to the Norway spruce reference genome *Picea abies* Nystedt et al. (2013)

Results

The

```
##   avgtWperSite avgtPperSite avgtajD
## 1   0.00317747  0.003939552 1.11307
```



$\theta = 4N_e\mu$

$D = \frac{\pi}{S}$

Table 1: Figure 1: Figure 2:

Conclusion

References

Nystedt, Björn, Nathaniel R. Street, Anna Wetterbom, Andrea Zuccolo, Yao Cheng Lin, Douglas G. Scofield, Francesco Vezzi, et al. 2013. "The Norway spruce genome sequence and conifer genome evolution." *Nature* 497 (7451). Nature Publishing Group: 579–84. https://doi.org/10.1038/nature12211.

Rigault, Philippe, Brian Boyle, Pierre Lepage, Janice E.K. Cooke, Jean Bousquet, and John J. MacKay. 2011. "A white spruce gene catalog for conifer genome analyses." *Plant Physiology* 157 (1): 14–28. https://doi.org/10.1104/pp.111.179663.

Yeaman, Sam, Kathryn A. Hodgins, Haktan Suren, Kristin A. Nurkowski, Loren H. Rieseberg, Jason A. Holliday, and Sally N. Aitken. 2014. "Conservation and divergence of gene expression plasticity following c. 140 million years of evolution in lodgepole pine (Pinus contorta) and interior spruce (Picea glauca × Picea engelmannii)." *New Phytologist* 203 (2): 578–91. https://doi.org/10.1111/nph.12819.