### Background

Trees in the genus *Populus* (poplars) are important for the bioenergy sector. Two closely related *Populus* species (*P.trichocarpa* and *P.balsamifera*) are of interest due to their ability to hybridize and adapt to local environments [1,2]. In the pacific northwest, there is overlap in these species geographic ranges which leads to mix ancestry of trees along this gradient. Such mixed ancestry may contribute to local adaptation and phenological response to environmental cues. To explore this relationship, stem cuttings were collected from 6 transcripts spanning the hybridization gradient. These cuttings were rooted and leaf tissue was collected for DNA extraction and paired end sequencing via the Illumina NovaSeq platform. Cuttings were transplanted to a common garden in Vermont where they were monitored for phenological changes. During the monitoring period, climate data was collected to assess the response of phenological processes to environment. I hypothesize there will be overlap in the genomic regions associated with phenological traits such as bud flush and response to climate characteristics. Further the genes shared among these traits are likely to be associated with breaking of dormancy at the cellular level.

### Bioinformatics pipeline

Chart, scatter chart

Description automatically generatedIn order to infer local ancestry, admixture was determined for all samples at K=5 and used to assign broad ancestry for each individual (balsamifera, trichocarpa, or admixed) based on the proportion of the genome attributed to each K group. Files containing *P.trichocarpa* and *P.balsamifera* reference individuals and the SNP positions were converted to vcf files with vcftools. For files containing admixed individuals, a while read loop was used read in files one line at a time and convert to vcf . These vcf files were matched to a sample ID and passed to loter which inferred local ancestry along chromosome 18. The text file output was converted from haploid format (0/1 for each homolog) to diploid (0/1/2 for both homologs combined) at each SNP site with datamash. SNP positions and two mock allele columns were added to the diploid file and a .fam file containing information about each sample was also generated. These files were converted to Plink format and used to determine local ancestry at each SNP site. Finally, Plink was used for admixture mapping by modeling the generalized linear relationship between local ancestry and phenotype or climate conditions. All Plink files were brought into R to analyze and visualize the relationship between climate or phenotype and proportion of *P.balsamifera* ancestry. The GenomicRanges package in R was used to identify outliers in admixture mapping results were used and overlap in outlier regions between all traits. Genes of interest were determined by aligning regions with reference genomes then matched to gene ontology (GO) terms using PopGenie.

***d***

***c***

***b***

***a***

***Figure 1.*** *The relationship of proportion of ‘P.balsamifera ancestry to a) chilling degree days, b) growing degree days, c) final freezing event, and d) bud flush. Each point represents a single individual collected along a transect.*

### Results

Timeline

Description automatically generatedProportion of *P.Balsamifera* ancestry showed a significant linear relationship with chilling degree days (CDD), growing degree days (GDD), final freezing event (FFE), and bud flush (figure 1). CDD, GDD, and FFE all showed positive relationships with ancestry where lower levels of *P.balsamifera* ancestry correlate to lower values for each trait. Bud flush, however, showed a signficant negative relationship with ancestry where lower levels of *P.balsamifera* ancestry correlate to higher values for budflush. Additionally, certain genomic regions show strong association with each trait (figure 2). The regions associated with these traits seem to be close but different for each trait. Further, when we look at inferred local ancestry along the chromosome, the regions of interest may be associated with changes in frequency of *P.trichocarpa* ancestry. Only GDD regions showed overlap with the bud flush regions, indicating that it may be the only climatic variable influencing bud flush. Unfortunately, when GO analysis was done on the overlapping regions there was no evidence of enrichment of currently classified functions. CDD, GDD, and FFE had 4, 6, and 4 candidate gene regions with 54, 49, and 27 genes respectively. GDD and FFE genes showed little enrichment for previously classified functions, where as CDD showed considerable function enrichment in GO categories “Biological Process” and “Molecular function”. Most of the enriched functions for CDD were associated with transporter activity.

***e***

***d***

***c***

***b***

***a***

***Figure 2.*** *The -log10 p-values at positions along the chromosome associated with a) chilling degree days, b) growing degree days, c) final freezing event, and d) bud flush. Red dots indicated outlier regions. Panel e represents inferred local ancestry along the chromosome.*

### Conclusion

Ancestry has an impact on both phenology and response to climatic conditions. Candidate gene regions for all traits were concentrated toward the end of the chromosome. Little overlap in gene regions was seen for climatic traits and bud flush, except for GDD. Additionally, there is enrichment among CDD genes for functions associated with transport which may be linked to the transport of biochemicals for winter storage or breaking of dormancy. It is worth noting that the lack of functional enrichment seen for GDD and FFE and the overlap of GDD with bud flush may be due to unknown or unclassified functions for candidate genes. Therefore, it would be naive to assume that candidate gene regions have no important function, especially considering that bud flush has previously covaried with important physiological traits across environmental gradients [2]. In future studies it would be helpful to pin down the roles of these genes in overall plant growth and phenology. Synthesizing the results, it is clear that admixed individuals may be more cold adapted as they are associated with a higher number of chilling degree day and a later final frost event. In future studies it would be useful to determine where specific gene regions were inherited from. On the surface, these results seem to indicate that regions of interest may be inherited from *P.trichocarpa* as indicated by high frequency of ancestry on chromosome 18, which is contrary to previous results [1].Lastly, it would be important expand the species used in the analysis to all North American poplars to understand how ancient introgression may impact modern phenotypes.

References

1. Suarez‐Gonzalez, Adriana, et al. “Scale and Direction of Adaptive Introgression between Black Cottonwood (Populus Trichocarpa) and Balsam Poplar (P. Balsamifera).” Molecular Ecology, vol. 27, no. 7, 2018, pp. 1667–1680.
2. Keller, Stephen R, et al. “Climate-Driven Local Adaptation of Ecophysiology and Phenology in Balsam Poplar, Populus Balsamifera L. (Salicaceae).” American Journal of Botany, vol. 98, no. 1, 2011, pp. 99–108.