

# Evaluating the Effects of Trembling Aspen Genetics on Insect Communities

Stat 877 project presentation

Clay Morrow and Ting-Fung Ma

May 2, 2019

# Background

## Plant-insect interactions:

- Plants are known to influence insect community structure
- Plant Diversity is positively correlated with insect diversity.
  - effects of interspecific variation are well studied
  - effects of **intraspecific variation** are less well studied

## *Populus tremuloides* (Trembling Aspen):

- the most wide-spread tree species in North America
- one of the fastest-growing tree species
- incredible amounts of intraspecific variation
- food source for an incredible diversity of herbivores

# Genome-wide association analysis (GWA)

Question: How do aspen genetics shape insect communities?

- Herbivorous insects are known to differ by aspen clone (genet) and by aspen traits (size, defense chemistry, nutrition, etc.).
- Is there evidence for genetic drivers other than these physical traits?
- **Are there genomic regions that affect multiple insects?**

GWA of insect incidence on aspen common garden:

- 18 common insect species; present/absent (1/0)
- 1,414 trees from 437 genets
- 8 tree trait covariates
- 4 time periods (longitudinal)
- 114,420 SNP regions with 3 possible genotypes {0,1,2}

# Statistical Model

$$\text{logit}(p_{ijk_g}) = \beta_0 + \alpha \mathbf{G}_{jg} + x_{jk_g}^T \beta + \varepsilon_{g(j)}$$

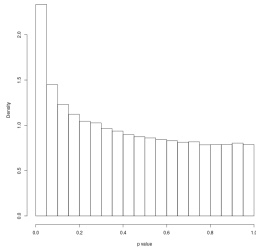
Where:

- $\text{logit}(y) = \log\left(\frac{y}{1-y}\right)$
- $p_{ijk_g}$ : probability of observing  $> 0$  individual insects of species  $i$  during survey event  $j$  on tree  $k$  of genet  $g$
- $\mathbf{G}_{jg}$ : SNP-specific genotype of genet  $g$
- $x_{jk_g}$ : vector of observed tree trait covariates for tree  $k$  during survey event  $j$
- $\varepsilon_{g(j)}$ : nested random effect of genet within survey event
- Storey's  $q$  values of coefficient  $\alpha$  used to select significant SNPs

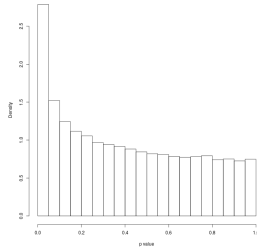
A total of 2,059,560 (18 insects  $\times$  114,420 SNPs) GLMMs were built.

# Diagnostics: p value distributions

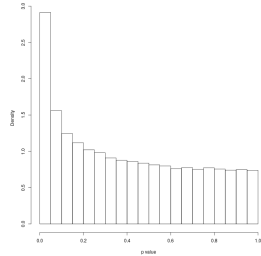
Histogram of Harmandia p values



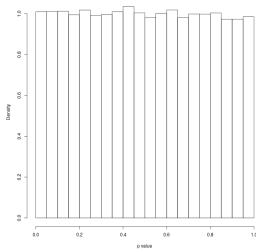
Histogram of Petiole Gall p values



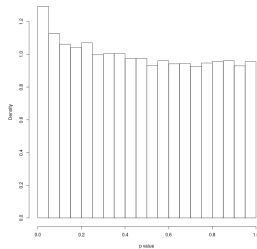
Histogram of Phyllocolpa p values



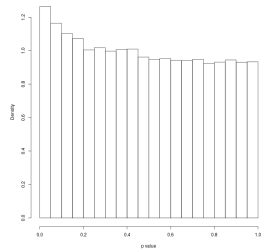
Histogram of Blotch Mine p values



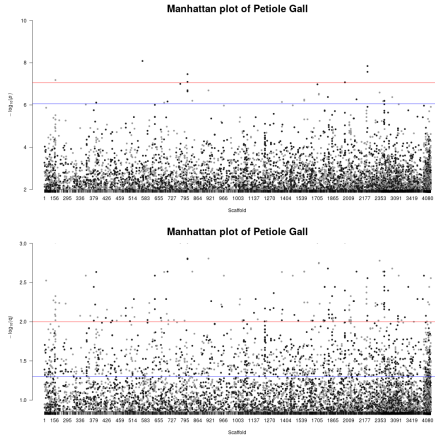
Histogram of Ants p values



Histogram of Green Sawfly p values



# Diagnostics: Bonferroni vs. Storey's q



- Using q values (bottom) allows for discovery of many more potential associations than bonferroni p values (top).
- significant q values contain all significant bonferroni p values

## Results: Significant SNPs per Insect

	insect description	SNPs	scaffolds
Green Aphids	free-feeding, specialist (salicaceae)	1788	1020
Petiole Gall	leaf-galling, specialist (populus)	1432	788
Phyllocolpa	leaf-rolling, specialist (salicaceae)	1363	814
Harmandia	leaf-galling, specialist (populus)	174	130
Smokey Aphids	free-feeding, specialist (populus)	90	64
Casebearer Moth	case-bearing, generalist	7	3
Lombardy Mine	leaf-mining, specialist (populus)	1	1
Cottonwood Leaf Mine	leaf-mining, specialist (salicaceae)	1	1
Leaf Edge Mine	leaf-mining, specialist (salicaceae)	0	0
Blotch Mine	leaf-mining, specialist (populus)	0	0
Weevil Mine	leaf-mining, specialist (salicaceae)	0	0
Blackmine Beetle	leaf-mining, specialist (populus)	0	0
Leafhoppers	free-feeding, generalist	0	0
Ants	aphid-tending, non-herbivore	0	0
Pale Green Notodontid	free-feeding, specialist (populus)	0	0
Aspen Leaf Beetle	free-feeding, specialist (populus)	0	0
Green Sawfly	free-feeding, specialist (populus)	0	0
Cotton Scale	scale insect, generalist	0	0

There are 4768 unique significant SNPs.

## Results: Common Associations

- Of these 4768 unique SNPs, There **84** SNPs with multiple insect associations.
- 4 of these had associations with 3 insects:

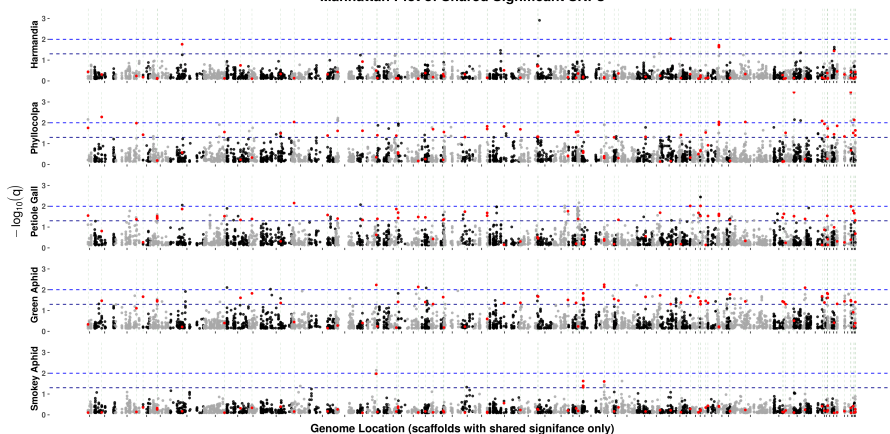
SNP.name	qval.Harmandia	qval.Phyllocolpa	qval.Petiole.Gall
Potra002191:26587	0.0193784	0.0123563	0.0286529
Potra002191:26642	0.0226734	0.0091460	0.0258001
Potra002191:26644	0.0226734	0.0091460	0.0258001
Potra002191:26650	0.0236877	0.0116280	0.0234547

- They are within 63 base pairs (likely in LD) which suggest there may be one important gene region.
- Note that these are the **2 galling insects** and one **leaf-rolling** (all stationary feeders, tightly associated with the leaves)



# Results: Multiple-Manhattan Plot

Manhattan Plot of Shared Significant SNPs



# Discussion

## Conclusions

- Evidence of genomic regions that influence individual insects and communities
- These genomic regions are independent of the observed tree traits

## Future Directions

- Impute gene annotations from congeneric *Populus trichocarpa* and model *Arabidopsis* to test for biological function/validation
- Compare with more traditional methods (using BLUPs for example)
  - A similar study on this garden found only 2 insect associations with  $< 12$  total significant SNPs using BLUPs and a more lenient inclusion cutoff ( $q < 0.1$ ) in 2015
- Include more tree traits and environmental variables