

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

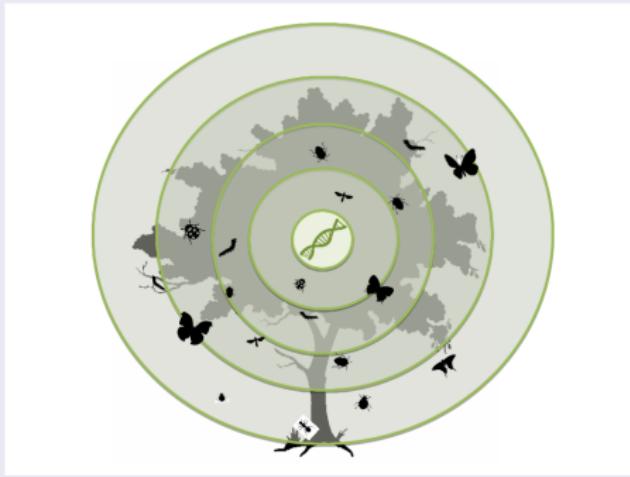
Common Garden



Research Goals

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?**



Genome-wide association analysis

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



Statistical Methods

Model: single marker associations

$$\text{logit}(p_{ijk_g}) = \beta_0 + \alpha G_g + x_{jk_g}^\top \beta + \varepsilon_{g(j)}$$

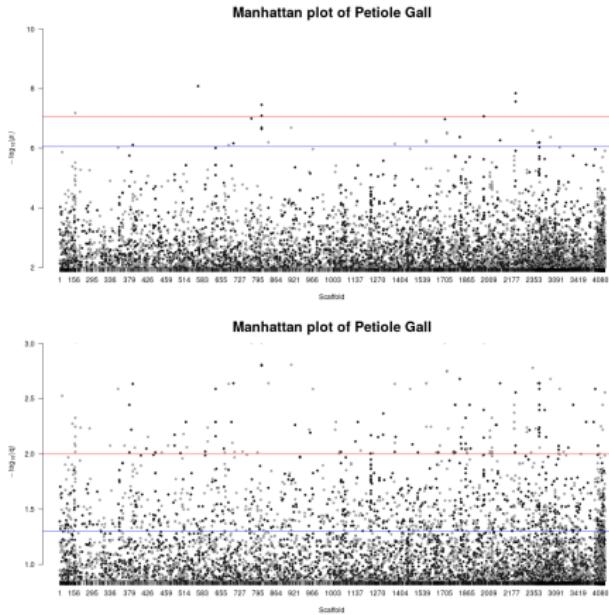
- $\text{logit}(y) = \log\left(\frac{y}{1-y}\right)$
- SNPs considered independently
- p_{ijk_g} : probability of observing > 0 individual insects of species i during survey event j on tree k of genet g
- G_g : SNP-specific genotype of genet g
- x_{jk_g} : vector of observed tree trait covariates for tree k during survey event j
- $\varepsilon_{g(j)}$: nested random effect of genet within survey event

Statistical Methods

Computation:

- `lme4::glmer()` is used to fit the GLMM with the nested random effect
- High Throughput Computing (CHTC) for parallelization
- A total of 2,059,560 (18 insects \times 114,420 SNPs) GLMMs were fit.
- Storey's q values of coefficient α used to select significant SNPs (.05 cutoff)

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

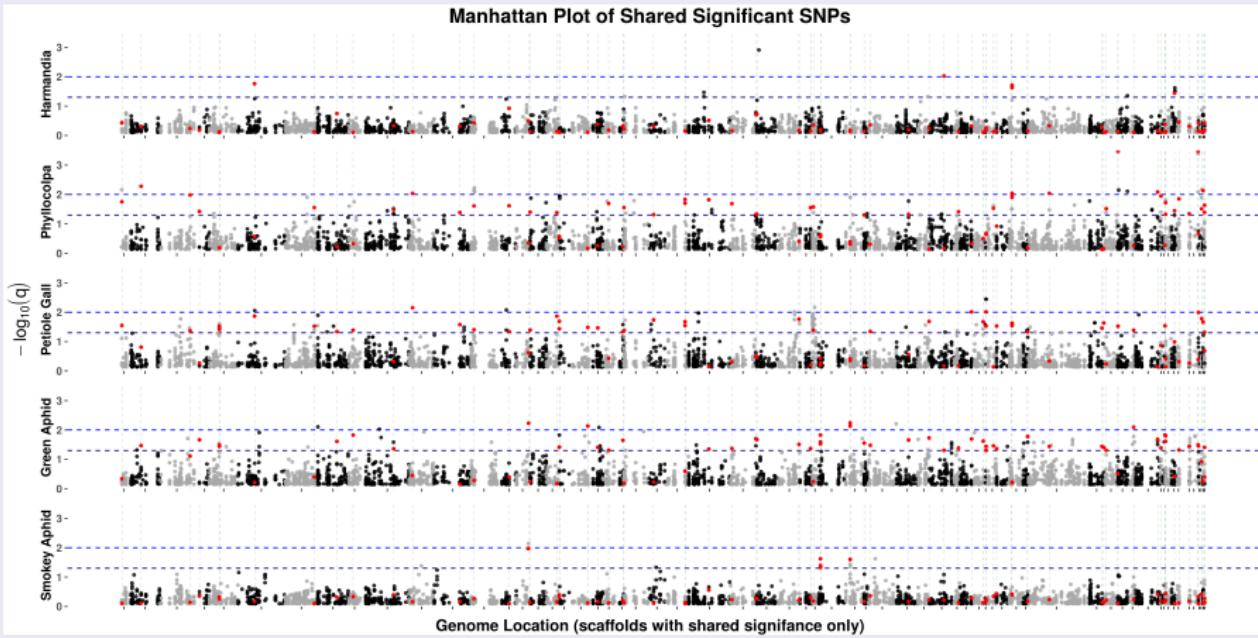
	SNPs	scaffolds	insect description
Green Aphids	1788	1020	free-feeding, specialist (salicaceae)
Petiole Gall	1432	788	leaf-galling, specialist (populus)
Phyllocolpa	1363	814	leaf-rolling, specialist (salicaceae)
Harmandia	174	130	leaf-galling, specialist (populus)
Smokey Aphids	90	64	free-feeding, specialist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

There are 4,768 unique significant SNPs.

Insects:



Results: Shared Associations



The corresponding genes of these 84 shared association markers (red) make up our candidate list.

Results: Imputed Gene Function

Top shared genes with annotations imputed from congeneric *Populus trichocarpa*:

Potri Gene	%match	Descr.	avg allele. freq	insects
Potri 006G267600.1	99.66	protein phosphatase 2C	0.110	Phyllocolpa, Petiole.Gall
Potri 013G036000.1	99.44	tRNA-specific adenosine deaminase	0.204	Phyllocolpa, Green.Aphids
Potri 007G058500.1	99.02	1-deoxy-D-xylulose-5-phosphate synthase	0.226	Phyllocolpa, Petiole.Gall
Potri 014G079900.1	99.01	dependent malic enzyme	0.112	Petiole.Gall, Green.Aphids
Potri 003G045300.1	99.01	multivesicular body protein	0.143	Phyllocolpa, Green.Aphids
Potri 011G157300.1	98.97	Uncharacterized oxidoreductase C663	0.136	Harmandia, Phyllocolpa, Petiole.Gall
Potri 006G249800.1	98.63	Probable boron transporter	0.139	Phyllocolpa, Petiole.Gall
Potri 016G142100.1	98.60	Unknown protein 1	0.150	Petiole.Gall, Green.Aphids
Potri 001G369000.1	98.59	Transmembrane emp24 domain-containing protein	0.205	Green.Aphids, Smokey.Aphids
Potri 016G037900.1	98.40	STRICTOSIDINE SYNTHASE-LIKE	0.170	Petiole.Gall, Green.Aphids
Potri 006G082900.1	98.38	Kinesin-like protein	0.218	Phyllocolpa, Green.Aphids
Potri 010G014100.1	98.33	repeat-containing protein	0.141	Petiole.Gall, Green.Aphids
Potri 015G112500.1	98.16	hydrolase domain-containing	0.126	Phyllocolpa, Petiole.Gall
Potri 014G156100.1	98.00	CSC1-like protein	0.137	Petiole.Gall, Green.Aphids
Potri 001G113500.1	98.00	domain-containing protein	0.403	Petiole.Gall, Green.Aphids
Potri 005G061600.1	97.99	glucuronate:xylan alpha	0.183	Petiole.Gall, Green.Aphids
Potri 002G009100.1	97.98	Bromodomain-containing protein	0.230	Green.Aphids, Smokey.Aphids
Potri 010G164400.1	97.93	uncharacterized protein LOC105116966 isoform X1	0.156	Petiole.Gall, Green.Aphids
Potri 006G149900.1	97.81	RNA-binding protein	0.150	Phyllocolpa, Green.Aphids
Potri 001G253900.1	97.81	Protein TIC 62, chloroplastic	0.222	Harmandia, Phyllocolpa
Potri 001G253800.1	97.67	Sugar transport protein	0.211	Harmandia, Phyllocolpa

Discussion

Conclusions

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

Future Directions

- More thoroughly research gene annotations for *Populus trichocarpa* and model *Arabidopsis thaliana*.
- Compare with other methods (multi-SNP, BLUP association, gBLUPs, etc)
 - A similar study on this garden found only 2 insect associations with < 12 total significant SNPs using BLUP association and a more lenient inclusion cutoff ($q < 0.1$) in 2015
- Include more tree traits and environmental variables and interactions
- Also run analyses with community diversity metrics as response variables.