Landscape genomics in *Populus trichocarpa*: disentangling the contributions of migration, hybridization and natural selection

Presenter: Quentin Cronk, UBC

Data sources

Plants:
* 498 *Populus trichocarpa* accessions from 30 populations ("drainages") ranging from Central Oregon, USA (44.3° N) to northern British Columbia, Canada (59.6° N)
* 10 accessions of the sister species *P. balsamifera*

SNPs:
34,131 SNPs in or near 3,543 genes from Illumina Infinium array (filtered, only loci with GenTrain score > 0.5, only genotypes with GenCall score > 0.15)


Individuality: one of the most important phenomena of life

- Individuals vary, and this variation has enormous significance for the functioning and distribution of ecosystems
- In poplar, our data suggest as an estimate that c. 95% of the variation between individuals can be accounted for by variation *within populations*
- And c. 5% of individual variation is accounted for *between populations* as geographically structured variation

Tackling the 5%: is geographical variation adaptive?

- Variation may result from introgression (and may be adaptive or neutral)
- May result from drift, i.e. population history (non-adaptive)
- May result from natural selection (adaptive)

Authors and contributors

- Armando Geraldes
- Nima Farzaneh
- Christopher J. Grassa
- Athena D. McKown
- Robert D. Guy
- Shawn D. Mansfield
- Carl J. Douglas
- Quentin C. B. Cronk

And POPCAN Team, UBC, UVic
(Genetic improvement of poplar trees as a Canadian bioenergy feedstock. Genome Canada/Genome BC, Project 168BIO; Carl Douglas and Shawn Mansfield, Project Leaders)
Population structure

Isolation by distance

Isolation by distance 2

Signatures of adaptation

- Exceptionally high Fst (Fdist2/Lositan and Bayescan)
- Significant correlation with climatic and geographical variables when population structure is taken into account (Bayenv)

Significant Fst outlier results (q≤0.05)

- Number of outliers: 1,007 SNPs in 603 genes
- Number assayed: 28,135 SNPs in 3,543 genes assayed
- 3.6% of snps are outliers
- 17% of genes contain outlier snps
What genes?

- Gene ontology (GO) analyses revealed that Fst outliers are statistically over-represented in certain gene classes
- Entire dataset: genes involved in circadian rhythm and response to red/far-red light
- Southern British Columbia: heat response genes are over-represented

Highest Fst values in dataset (>0.35)

- FAR1 (FAR-RED IMPAIRED RESPONSE 1); transcription activator/ transcription factor
- AtRABA4a (Arabidopsis Rab GTPase homolog A4a); GTP binding
- ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family
- LUG (LEUNIG); protein binding / protein heterodimerization/ transcription repressor

Genes with most significant correlation with climate/geographical factors (log10(BF)>5)

<table>
<thead>
<tr>
<th>GENE</th>
<th>CLIMATE VARIABLE</th>
<th>Cor./BF</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBP-18 (DNA binding, transcription factor)</td>
<td>Continentality (mean temp. warmest month-coldest month)</td>
<td>5.290</td>
</tr>
<tr>
<td>LUH (LEUNIG HOMOLOG)</td>
<td>Mean annual temperature</td>
<td>5.035</td>
</tr>
<tr>
<td>REV (REVOLUTA)</td>
<td>Latitude</td>
<td>5.118</td>
</tr>
<tr>
<td>PRR7 (PSEUDO-RESPONSE REGULATOR 7)</td>
<td>Latitude, Extreme max. temp., Evaporation</td>
<td>5.886, 5.038, 5.910</td>
</tr>
<tr>
<td>O-fucosyltransferase family protein</td>
<td>Chilling degree days (days below zero)</td>
<td>5.531</td>
</tr>
<tr>
<td>TRANSPARENT TESTA GLABRA 1</td>
<td>Continentality</td>
<td>6.067</td>
</tr>
<tr>
<td>unknown protein; functional in chloroplast</td>
<td>Continentality</td>
<td>6.211</td>
</tr>
</tbody>
</table>

How much of the “5%” is a result of adaptive evolution, how much population history?

- Difficult to say, but best guess is 50:50 (perhaps 2.5% of total genetic variance of poplar driven by local adaptation?)
- However, what does emerge without doubt is a picture of pervasive adaptation to geography and climate over the range of P. trichocarpa