Using genomics to study tanoak’s past, present, and future

KJ Hayden, M Garbelotto, R Cronn, BJ Knaus, H Rai, JW Wright
Current Knowledge

• Partition of neutral genetic diversity
  – Range-wide

• Contribution of clonality to stand genetic diversity
  – Within a stand

• Pollination and outcrossing to parentage
  – Within a tree

• Contribution of gene expression to phenotype
  – Within a genotype
DNA \rightarrow RNA \rightarrow Proteins \rightarrow Metabolites \rightarrow Ions
The present

First experiment
• One tree genotype
• One pathogen genotype
• One mock-inoculation control
• 2 time intervals from inoculation
5.2 Gbp, in 64 M 80bp fragments
Align to the *P. ramorum* reference to subtract pathogen sequences

Reference genome: Tyler et al. 2006.
Science 313: 1261-66
Align to the *P. ramorum* reference to subtract pathogen sequences

<table>
<thead>
<tr>
<th></th>
<th>% match <em>P. ramorum</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 day inoculated</td>
<td>1.2%</td>
</tr>
<tr>
<td>1 day control</td>
<td>1.4%</td>
</tr>
<tr>
<td>5 day inoculated</td>
<td>4.5%</td>
</tr>
<tr>
<td>5 day control</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

5 day inoculated

<table>
<thead>
<tr>
<th>Pathogen sequence type</th>
<th>Transcripts matched</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. ramorum</em> and <em>P. sojae</em> avirulence homologs (Jiang et al. 2008)</td>
<td>41</td>
</tr>
<tr>
<td><em>P. ramorum</em> cellulose-binding, elicitin, lectin-like, and necrosis-inducing proteins (assembled from Tyler et al. 2006)</td>
<td>63</td>
</tr>
<tr>
<td><em>P. infestans</em> CBEL, CRN, and necrosis-suppressing genes (Sierra et al. 2010)</td>
<td>16</td>
</tr>
<tr>
<td>Additional BLAST annotation</td>
<td>1893</td>
</tr>
</tbody>
</table>
What to do when there is no reference?

• As sequencing technology has improved, so have methods for analyzing the data
• Reads that did not map to *P. ramorum* pooled and assembled together into a de novo reference
  • 48,388 sequences
  • 34,642 unique isogroups
  • N50 = 1289
• Individual lanes mapped back to reference to count expression

(Trinity, Grabherr et al. 2011 Nature Biotechnology)
Differential expression, 5 days post-inoculation

844 genes up- or down-regulated
• 102 match disease-related genes in Chestnut (A. Barakat, NC State)
• 47 R-genes
• Others annotated from published genomes, NCBI databases, and GO annotations
The future

We have

• Pipeline for inoculation, sequencing, and analysis
• The first tanoak transcriptome reference
• A first look at what is different in infected tissue vs not infected
• Lots of questions unanswered-
  – Variation among trees?
    • Only 1 genotype sampled
  – Variation among pathogen genotypes?
    • Only 1 genotype sampled
  – Confirmed function of DE genes
The future

• Expand sequencing
  – Resistant and susceptible sibs from across populations
    • Integrate topics discussed here – elicitors, phenolics, transposons
  – Phosphonate-treated vs control
    • Systemic fungicide, with direct and indirect action
    • Mechanism not well-understood; may involve increases in pathogen elicitors (Grant et al. 1990)

• Reference and raw data will be freely available (dendrome)
  – For tanoak gene expression
  – For comparative genomics (better reconstruct tanoak’s past)
  – For other queries
    • sequence is meaningless if we don’t know what it does – each addition to the database adds to our understanding of biology
Acknowledgements

• USDA Forest Service, Pacific Southwest Research Station
• USDA Forest Service, Pacific Northwest Research Station
• USDA Forest Service, Rocky Mountain Research Station
• The Gordon and Betty Moore Foundation
• Eriko Miura, Tara Jennings, Kenan Celtik