Benchmarking BAli-Phy
Experiments

• Testing multiple sequence alignment and tree generation from BAli-Phy vs. two other methods to understand:
  – Alignment Error
  – Tree Error
  – Convergence Rates / Running Time

• 9 Simulated Datasets (Actually 8):
  – 12, 25, 100 taxa
  – 3 Rates of Evolution (named M2, M3, M4 in order from hardest to easiest)
    – M3-12 had an error 😞

• 4 Alignment, 3 Tree Methods:
  – BAli-Phy MAP alignment/tree
  – BAli-Phy Posterior Decoding alignment only
  – PASTA alignment/tree
  – MAFFT alignment + RAxML tree
Results
Relative Results (Note: All Axes Vary)

Notes:
- All data sets have very low alignment error except for M2-100 (the hardest).
- Tree error rates are quite high. Possibly an artifact of Indelible data.
- Definitions:
  - SPFN: False Negative Rate
  - SPFP: False Positive Rate
  - TC: Total Column Score
  - Rfnorm: Tree Error (RF Distance Normalized)
Results

Absolute Results

Method
- BP - MAP
- BP - Posterior Decoding
- MAFFT / RAxML
- PASTA

SPFN

Size=12
Size=25
Size=100

M2
M3
M4

Rate of Evolution

SPFP

Size=12
Size=25
Size=100

M2
M3
M4

Rate of Evolution

TC

Size=12
Size=25
Size=100

M2
M3
M4

Rate of Evolution

Method

BP - MAP
BP - Posterior Decoding
MAFFT / RAxML
PASTA
## Results

**BAli-Phy Convergence (ESS) Values by Parameter**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Size 12</th>
<th>Size 25</th>
<th>Size 100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M2</td>
<td>M3</td>
<td>M4</td>
</tr>
<tr>
<td>prior</td>
<td>1,499</td>
<td>2,987</td>
<td>1,338</td>
</tr>
<tr>
<td>prior_A1</td>
<td>66,472</td>
<td>42,900</td>
<td>19,993</td>
</tr>
<tr>
<td>likelihood</td>
<td>90,049</td>
<td>73,589</td>
<td>41,533</td>
</tr>
<tr>
<td>logp</td>
<td>2,062</td>
<td>2,987</td>
<td>1,850</td>
</tr>
<tr>
<td>Main.mu1</td>
<td>417,929</td>
<td>413,302</td>
<td>176,506</td>
</tr>
<tr>
<td>S1.TN.kappaPur</td>
<td>160,939</td>
<td>163,454</td>
<td>56,436</td>
</tr>
<tr>
<td>S1.TN.kappaPyr</td>
<td>171,642</td>
<td>177,824</td>
<td>48,557</td>
</tr>
<tr>
<td>S1.F.piA</td>
<td>103,396</td>
<td>101,917</td>
<td>26,900</td>
</tr>
<tr>
<td>S1.F.piG</td>
<td>95,433</td>
<td>94,631</td>
<td>35,145</td>
</tr>
<tr>
<td>S1.F.piT</td>
<td>101,148</td>
<td>100,422</td>
<td>36,201</td>
</tr>
<tr>
<td>S1.F.piC</td>
<td>94,158</td>
<td>92,904</td>
<td>31,277</td>
</tr>
<tr>
<td>I1.RS07.logLambda</td>
<td>396,675</td>
<td>313,605</td>
<td>165,596</td>
</tr>
<tr>
<td>I1.RS07.meanIndelLength</td>
<td>204,421</td>
<td>202,094</td>
<td>98,962</td>
</tr>
<tr>
<td></td>
<td>A1</td>
<td></td>
<td>47,739</td>
</tr>
<tr>
<td>#indels1</td>
<td>55,478</td>
<td>37,838</td>
<td>17,013</td>
</tr>
<tr>
<td></td>
<td>indels1</td>
<td></td>
<td>47,735</td>
</tr>
<tr>
<td>#substs1</td>
<td>125,044</td>
<td>124,461</td>
<td>27,980</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td>349,651</td>
</tr>
</tbody>
</table>