<table>
<thead>
<tr>
<th>Dataset</th>
<th>Million reads</th>
<th>Kmer</th>
<th>No. of Contigs</th>
<th>Largest contig (nt)</th>
<th>Complexity (nt)</th>
<th>N50</th>
<th>MAC genome coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>126</td>
<td>45</td>
<td>38714</td>
<td>159958</td>
<td>96838332</td>
<td>18877</td>
<td>96.29%</td>
</tr>
<tr>
<td>1</td>
<td>126</td>
<td>51</td>
<td>32891</td>
<td>177346</td>
<td>97451433</td>
<td>26902</td>
<td>96.38%</td>
</tr>
<tr>
<td>1</td>
<td>126</td>
<td>55</td>
<td>42969</td>
<td>82851</td>
<td>99085030</td>
<td>7298</td>
<td>97.30%</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>45</td>
<td>94491</td>
<td>107860</td>
<td>61535444</td>
<td>1087</td>
<td>49.74%</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>51</td>
<td>80141</td>
<td>86502</td>
<td>59448180</td>
<td>1180</td>
<td>46.45%</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>55</td>
<td>71678</td>
<td>82851</td>
<td>58477670</td>
<td>1230</td>
<td>44.89%</td>
</tr>
<tr>
<td>3</td>
<td>121</td>
<td>45</td>
<td>26721</td>
<td>158845</td>
<td>95582645</td>
<td>24907</td>
<td>95.91%</td>
</tr>
<tr>
<td>3</td>
<td>121</td>
<td>51</td>
<td>23091</td>
<td>187719</td>
<td>96119509</td>
<td>29755</td>
<td>95.65%</td>
</tr>
<tr>
<td>3</td>
<td>121</td>
<td>55</td>
<td>22059</td>
<td>138361</td>
<td>96437319</td>
<td>25160</td>
<td>96.63%</td>
</tr>
<tr>
<td>4</td>
<td>121</td>
<td>55</td>
<td>20728</td>
<td>243532</td>
<td>96552463</td>
<td>32589</td>
<td>96.42%</td>
</tr>
</tbody>
</table>

Table S1. Assembly statistics. The datasets are as follows: 1) All PGM reads; 2) All pairs of PGM reads with at least one read that does not map to the MAC reference genome; 3) All PGM reads after removal of those reads that do not match a putative MAC IES junction identified by the MIRAA pipeline; 4) All PGM reads after removal of those reads that do not match a MAC IES junction identified using the MICA pipeline and the first 9 assemblies. Paired-end assembly was carried out using the Velvet short read assembler (version 1.0.18) with the indicated Kmer value. The MAC genome coverage by the assembled contigs was determined by mapping them to the reference genome with BLAT.