Additional material

Mobilomics in *S. cerevisiae* strains

Giulia Menconi\textsuperscript{1}, Giovanni Battaglia\textsuperscript{2},
Roberto Grossi\textsuperscript{2}, Nadia Pisanti\textsuperscript{2} and Roberto Marangoni\textsuperscript{2,3}

\textsuperscript{1}Istituto Nazionale di Alta Matematica, Città Universitaria, 00185 Roma, Italia
\textsuperscript{2}Dipartimento di Informatica, Università di Pisa, 56127 Pisa, Italia.
\textsuperscript{3}CNR-Istituto di Biofisica, 56124 Pisa, Italia.

\textbf{REGENDER performance and complexity}

The evaluation of \textsc{regender} performance and complexity has been discussed in the Master thesis “Strumenti computazionali per la mobilomica, una nuova branca della bioinformatica” by Emiliano Biscardi, presented at the University of Pisa, in March 2011, and publicly available at the URL: http://etd.adm.unipi.it/theses/available/etd-02202011-152231. In this work, Biscardi compared \textsc{regender} with several of the most commonly used alignment tools. A quantitative comparison concerning the space and time complexity has been measured on the given dataset. A qualitative comparison concerning the location and the length of the conserved sequences detected by each tool has been also performed. Both kinds of comparison are summarized in the enclosed table. The comparison of the conserved regions detected is shown in the enclosed figure, where the case of ChrIV of \textsc{RefSeq} and \textsc{Y55} is presented. The experimented tools were: \textsc{AVID} \cite{2}, \textsc{BLAST}\cite{1}, \textsc{GSALIGN}\cite{10}, \textsc{LAGAN} \cite{3}, \textsc{LASTZ} \cite{6}, \textsc{MGA} \cite{7}, \textsc{MUMMER} \cite{4,5,8}, \textsc{MURASAKI} \cite{9}.

\textbf{References}

\begin{enumerate}
\end{enumerate}
<table>
<thead>
<tr>
<th>Tool</th>
<th>Time</th>
<th>Resident memory</th>
<th>Conserved regions</th>
<th>Shared coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVID</td>
<td>+ 764.09 %</td>
<td>+ 54.88 %</td>
<td>+ 0.99 %</td>
<td>98.64 %</td>
</tr>
<tr>
<td>BLAST</td>
<td>- 60.94 %</td>
<td>- 89.29 %</td>
<td>- 2.18 %</td>
<td>96.4 %</td>
</tr>
<tr>
<td>GSALIGN</td>
<td>+ 477.90 %</td>
<td>- 47.12 %</td>
<td>- 17.74 %</td>
<td>80.37 %</td>
</tr>
<tr>
<td>LAGAN</td>
<td>+ 995.30 %</td>
<td>+ 272.15 %</td>
<td>+ 0.87 %</td>
<td>98.30 %</td>
</tr>
<tr>
<td>LASTZ</td>
<td>+ 244.75 %</td>
<td>- 56.04 %</td>
<td>- 15.74 %</td>
<td>82.54 %</td>
</tr>
<tr>
<td>MGA</td>
<td>- 2.76 %</td>
<td>- 94.06 %</td>
<td>- 64.97 %</td>
<td>34.82 %</td>
</tr>
<tr>
<td>MUMMER</td>
<td>- 56.63 %</td>
<td>- 91.20 %</td>
<td>- 17.69 %</td>
<td>81.37 %</td>
</tr>
<tr>
<td>MURASAKI</td>
<td>+ 275.41 %</td>
<td>+ 1035.87 %</td>
<td>- 25.59 %</td>
<td>70.12 %</td>
</tr>
</tbody>
</table>

Table 1: Comparison between REGENER and the most used alignment tools. For each tool, it is shown the percentage of the average variation of the following parameters, when compared to the corresponding ones of REGENER: execution time, allocated resident memory, conserved regions found, shared aligned bases.


Figure 1: **Comparison of identified conserved regions.** All the plots refer to ChrIV in RefSeq (top) and in Y55 (bottom). The shadowed regions are detected as conserved by the tool indicated in the top line.