Model selection emphasises the importance of non-chromosomal information in genetic studies

Reda Rawi\textsuperscript{1}, Mohamed El Anbari\textsuperscript{1,2}, Halima Bensmail\textsuperscript{1,*}

\textsuperscript{1} Computational Science and Engineering Center, Qatar Computing Research Institute, Doha, Qatar
\textsuperscript{2} Division of Biomedical Informatics, Sidra Medical and Research Center, Doha, Qatar
* E-mail: hbensmail@qf.org.qa

Table S3: Frequency of predictors ($X_1$, $X_2$, and $X_1X_2$) within 1000 modelling repeats using LASSO for hierarchical interactions.

<table>
<thead>
<tr>
<th>Gene deletion</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_1X_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCM22</td>
<td>57</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>PEP12</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>PEP7</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>PHO88</td>
<td>640</td>
<td>626</td>
<td>467</td>
</tr>
<tr>
<td>SKI8</td>
<td>1000</td>
<td>997</td>
<td>997</td>
</tr>
<tr>
<td>VPS16</td>
<td>1000</td>
<td>874</td>
<td>870</td>
</tr>
<tr>
<td>PHO88(killer)</td>
<td>1000</td>
<td>956</td>
<td>953</td>
</tr>
<tr>
<td>PHO88(non-killer)</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>SKI8(killer)</td>
<td>1000</td>
<td>988</td>
<td>988</td>
</tr>
<tr>
<td>SKI8(non-killer)</td>
<td>998</td>
<td>995</td>
<td>995</td>
</tr>
</tbody>
</table>

Aside from the MCM22 and PHO88(non-killer) gene deletion experiments, all variables ($X_1$, $X_2$, and $X_1X_2$) are selected in most modelling repeats. An exception is the PHO88 deletion experiment where the interaction predictor ($X_1X_2$) is selected in only half of the 1000 repeats.