Journal Club –
Yeast Papers from Kruglyak Lab
8/13/2013

Han Fang
hfang@cshl.edu
Lyon Lab
Contents

• Quantitative trait locus (QTL) analysis (2008)
• Finding the sources of missing heritability in a yeast cross (2013)
• Genetics of single-cell protein abundance variation in large yeast populations (2013, In press) To discuss next time
Quantitative trait locus (QTL) analysis

Some major questions and techniques

• A few loci with large effects or many loci with minute effects?
• Sample size matters? Interactions?
• Limitations: large sample size, only map differences between the initial parental strains, specific alleles may not be relevant to natural populations, small number of genes were identified
• eQTL, pQTL
• Maybe in combination with GWAS or combination of different QTL

Finding the sources of missing heritability in a yeast cross

Motivation

• GWAS has underscored the problem of missing heritability.

• Sample sizes, minimal effects to ever be individually detected.

• Epistasis interactions may inflate heritability measures.

• Structural variations, G x E interactions, parent of origin effects, heritable epigenetic factors

• Direct estimates of heritability, G x G interactions, or locus effect size

Methods

• Construction of segregant panel (1008)
• Sequencing step
• Phenotyping by end-point growth on agar plates
• Definition of genetic factors
• Calculating heritability
• QTL mapping

The design of the segregant panel

Methods

• Construction of segregant panel (1008)
• **Sequencing step**
• Phenotyping by end-point growth on agar plates
• Definition of genetic factors
• Calculating heritability
• QTL mapping

Sequencing step

• DNA preparation and sequencing library construction
• Illumia HiSeq 2000 sequencing
• sequenced the parent strains to high coverage
• compared the sequences to define 30,594 high-confidence SNPs

Counts of sequencing reads at SNP sites are plotted (Y-axis) against genome position (X-axis) for a representative segregant; the orange (BY) and purple (RM) bars indicate parental haplotype calls, and the vertical black bars delineate chromosomes.

Methods

- Construction of segregant panel (1008)
- Sequencing step
- Phenotyping by end-point growth on agar plates
- Definition of genetic factors
- Calculating heritability
- QTL mapping

Phenotyping by end-point growth

- Custom R code was written to determine the size of each colony.
- Images were segmented using k-means clustering on the distribution of pixel intensities across a plate.
- Radius: $\sqrt{\frac{\text{pixelcount}}{\pi}}$.
- QC with certain thresholds
- A robust locally weighted regression was fit to the radius measurements

(C) An image of endpoint colony growth is shown for 384 segregants, with the outlines of colonies, as detected by our image processing software, indicated in red.

Spearman correlation coefficients for all pairs of traits are shown. Numbers in table cells indicate (100 * correlation coefficient).

Methods

• Construction of segregant panel (1008)
• Sequencing step
• Phenotyping by end-point growth on agar plates
• Definition of genetic factors
• Calculating heritability
• QTL mapping

Definition of genetic factors

• Phenotypic variation:
  contribution of heritable genetic factors (broad-sense heritability)
  measurement errors
  other random environmental effects.
• Broad-sense heritability:
  contribution of additive genetic factors (narrow-sense heritability)
  dominance effects
  gene–gene interactions (differences)
  gene–environment interactions

Methods

• Construction of segregant panel (1008)
• Sequencing step
• Phenotyping by end-point growth on agar plates
• Definition of genetic factors
• Calculating heritability
• QTL mapping

Calculating heritability

- Broad-sense heritability was calculated using replicated segregant data and a random effects analysis of variance.
\[
\frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2}
\]

- Narrow-sense heritability was calculated for each trait using a linear mixed model.
\[
y = \beta 1_n + Zu + e \quad \rightarrow \quad V = A\sigma_A^2 + I\sigma_{EY}^2 \quad \rightarrow \quad \frac{\sigma_A^2}{\sigma_A^2 + \sigma_{EY}^2}
\]

Heritability for 46 yeast traits.

Methods

• Construction of segregant panel (1008)
• Sequencing step
• Phenotyping by end-point growth on agar plates
• Definition of genetic factors
• Calculating heritability
• QTL mapping

QTL mapping

• a step-wise forward-search approach to detect QTL
• build a multiple-regression model
• detected a total of 591 QTL for 46 traits at an empirical false-discovery rate (FDR) of 5%
• observed varying degrees of trait complexity, with a minimum of 5, a maximum of 29 and a median of 12 QTL per trait

A histogram of QTL effect sizes across all traits is plotted, showing that most detected QTL have small effects. Effect size here is the absolute value of the standardized difference in allelic means for each QTL. (The blue line : a fit of a truncated exponential distribution of effect sizes.)

For each trait, the effect sizes of detected QTL are sorted from largest to smallest, and the cumulative phenotypic variance explained is plotted (Y-axis) against the number of detected QTL (X-axis).
Most additive heritability is explained by detected QTL.

QTL detection for a complex trait.

(B) Statistical power are shown for mapping populations of 100 (blue) and 1000 (red) segregants at a genomewide significance threshold.

Prediction of segregant trait values from QTL phenotypes.

QTL-QTL interactions

• differences between the estimates of broad-sense and narrow-sense heritability.
• two-locus interactions.
• first performed an exhaustive two-dimensional scan for pairwise interactions (large search space, low power)
• 17 of the 46 traits, with a total of 23 interacting locus pairs
QTL-QTL interactions – cont.

- testing only for interactions between each locus with significant additive effects and the rest of the genome
- detected interactions for 24 of the 46 traits, with a total of 78 QTL–QTL interactions
- a minimum of 1 and a maximum of 16 pairwise interactions per trait
Non-additive genetic variance explained by QTL–QTL interactions.

Conclusion

• Large panel of segregants from a cross in two yeast strains, short read sequencing
• total and additive heritability, and interactions
• Consistent with the suggestions that missing additive heritability arises primarily from many loci with small but not infinitesimal effects. (Can be discovered with large sample sizes)
• Future work: delineate the contributions of common and rare variants to inherited variation
Genetics of single-cell protein abundance variation in large yeast populations

Table 1 – mRNA-specific and protein-specific local QTL

<table>
<thead>
<tr>
<th>Gene</th>
<th>X-pQTL LOD</th>
<th>eQTL LOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local eQTL only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YJL201W</td>
<td>0.5</td>
<td>15.2</td>
</tr>
<tr>
<td>YPL048W</td>
<td>0.4</td>
<td>7.3</td>
</tr>
<tr>
<td>YDL171C</td>
<td>0.5</td>
<td>6.4</td>
</tr>
<tr>
<td>YLR438W</td>
<td>1.0</td>
<td>6.4</td>
</tr>
<tr>
<td>YNL044W</td>
<td>0.5</td>
<td>5.3</td>
</tr>
<tr>
<td>Local X-pQTL only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YJL130C</td>
<td>6.4</td>
<td>0.2</td>
</tr>
<tr>
<td>YDL126C</td>
<td>13.7</td>
<td>0.2</td>
</tr>
<tr>
<td>YGL026C</td>
<td>8.6</td>
<td>0.1</td>
</tr>
<tr>
<td>YMR315W</td>
<td>12.7</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Albert et.al. In presss
<table>
<thead>
<tr>
<th>chromosome</th>
<th>Position (peak SNP)</th>
<th>% of genes regulated at LOD &gt; 4.5 / LOD &gt; 3</th>
<th>mRNA hotspot¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>39,010</td>
<td>31 / 40</td>
<td>Glu1</td>
</tr>
<tr>
<td>II</td>
<td>132,948</td>
<td>31 / 41</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>397,978</td>
<td>9 / 18</td>
<td>Glu2</td>
</tr>
<tr>
<td>IV</td>
<td>223,943</td>
<td>12 / 24</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>192,064</td>
<td>16 / 31</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>371,845</td>
<td>16 / 21</td>
<td>Glu6</td>
</tr>
<tr>
<td>VII</td>
<td>137,332</td>
<td>15 / 26</td>
<td>-</td>
</tr>
<tr>
<td>VII</td>
<td>505,871</td>
<td>16 / 29</td>
<td>-</td>
</tr>
<tr>
<td>VIII</td>
<td>103,041</td>
<td>19 / 29</td>
<td>Glu7</td>
</tr>
<tr>
<td>VIII</td>
<td>419,747</td>
<td>8 / 12</td>
<td>-</td>
</tr>
<tr>
<td>X</td>
<td>142,009</td>
<td>18 / 26</td>
<td>-</td>
</tr>
<tr>
<td>X</td>
<td>655,465</td>
<td>11 / 15</td>
<td>-</td>
</tr>
<tr>
<td>XI</td>
<td>234,462</td>
<td>16 / 23</td>
<td>Glu8</td>
</tr>
<tr>
<td>XII</td>
<td>238,302</td>
<td>16 / 31</td>
<td>-</td>
</tr>
<tr>
<td>XII</td>
<td>656,893</td>
<td>41 / 49</td>
<td>Glu9</td>
</tr>
<tr>
<td>XII</td>
<td>1,039,502</td>
<td>12 / 19</td>
<td>Yvert¹</td>
</tr>
<tr>
<td>XIII</td>
<td>96,832</td>
<td>31 / 46</td>
<td>Glu10</td>
</tr>
<tr>
<td>XIV</td>
<td>232,509</td>
<td>13 / 19</td>
<td>-</td>
</tr>
<tr>
<td>XIV</td>
<td>465,007</td>
<td>58 / 65</td>
<td>Glu11</td>
</tr>
<tr>
<td>XV</td>
<td>162,766</td>
<td>56 / 70</td>
<td>Glu12</td>
</tr>
</tbody>
</table>

¹ As identified in Smith & Kruglyak 2008¹¹.

² This hotspot was not observed in Smith & Kruglyak¹¹, but was present in an earlier BY/RM eQTL dataset⁴⁷.
<table>
<thead>
<tr>
<th>Gene</th>
<th>chrXI effect</th>
<th>chrXII effect</th>
<th>Chr XV effect</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP14*</td>
<td>-0.35</td>
<td>-0.14</td>
<td>-0.14</td>
<td>ATP synthase</td>
</tr>
<tr>
<td>ATP17*</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.18</td>
<td>ATP synthase</td>
</tr>
<tr>
<td>ATP2*</td>
<td>-0.21</td>
<td>-0.3</td>
<td>-0.22</td>
<td>ATP synthase</td>
</tr>
<tr>
<td>CIT1*</td>
<td>-0.23</td>
<td>-0.36</td>
<td>-0.26</td>
<td>Citrate synthase</td>
</tr>
<tr>
<td>MDH1*</td>
<td>-0.22</td>
<td>-0.1</td>
<td>-0.39</td>
<td>Malate Dehydrogenase</td>
</tr>
<tr>
<td>ADO1</td>
<td>-0.09</td>
<td>-0.25</td>
<td>0.09</td>
<td>Adenosine kinase</td>
</tr>
<tr>
<td>GLTI</td>
<td>-0.08</td>
<td>0.13</td>
<td>0.24</td>
<td>Glutamate synthase</td>
</tr>
<tr>
<td>LIA1</td>
<td>-0.1</td>
<td>0.15</td>
<td>0.15</td>
<td>Deoxyhypusine hydroxylase</td>
</tr>
<tr>
<td>TDH3</td>
<td>-0.14</td>
<td>0.35</td>
<td>0.27</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)</td>
</tr>
<tr>
<td>YHB1</td>
<td>-0.16</td>
<td>-0.92</td>
<td>0.13</td>
<td>Nitric oxide oxidoreductase</td>
</tr>
<tr>
<td>YLR179C</td>
<td>-0.09</td>
<td>0.7</td>
<td>0.17</td>
<td>Unknown function</td>
</tr>
</tbody>
</table>

* involved in aerobic respiration

Albert et.al. In presss
Figure 1 – Distant and local variation affects protein levels

Albert et.al. In presss
Figure 2 – X-pQTL hotspots and overlap with loci affecting mRNA abundance

Albert et al. In press
Figure 3 – Hotspot effects
Albert et.al. In presss
Fig. S1 Overview of the experimental design
Albert et.al. In presss
Supplementary Figure S2 – Illustration of FACS design
Supplementary Figure S3 – Sequence analyses and X-pQTL detection example

Albert et al. In press
Supplementary Figure S4 – Reproducibility examples
Albert et.al. In presss
Supplementary Figure S5 – The impact of small effect sizes on the π1 estimate
Albert et.al. In press
Supplementary Figure S6 – Genes regulated by the hotspots on chromosomes XI, XII, and XV involved in aerobic respiration

Albert et.al. In presss
References

