# 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

Miles, C. & Wayne, M. (2008) Quantitative trait locus (QTL) analysis. Nature Education 1(1)
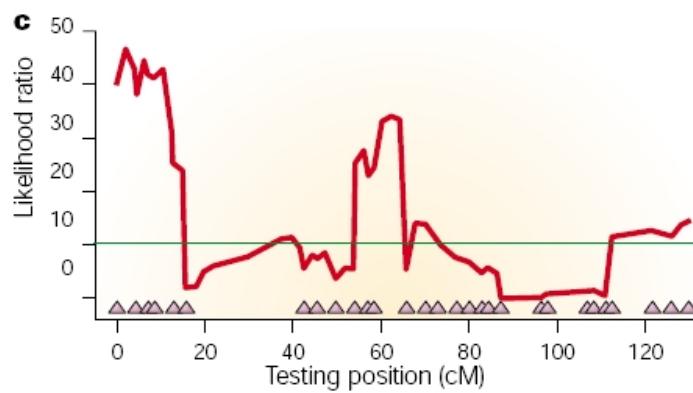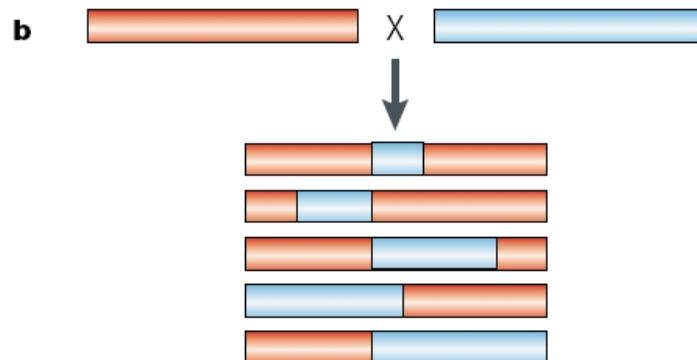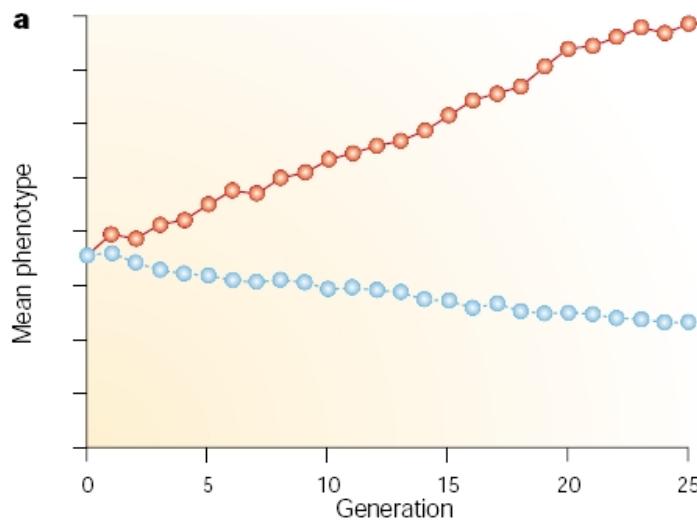
# Some major questions and techniques

- 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

Miles, C. & Wayne, M. (2008) Quantitative trait locus (QTL) analysis. Nature Education 1(1)

# Finding the sources of missing heritability in a yeast cross

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# 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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# Methods

- **Construction of segregant panel (1008)**

- Sequencing step

- Phenotyping by end-point growth on agar plates

- Definition of genetic factors

- Calculating heritability

- QTL mapping

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# The design of the segregant panel
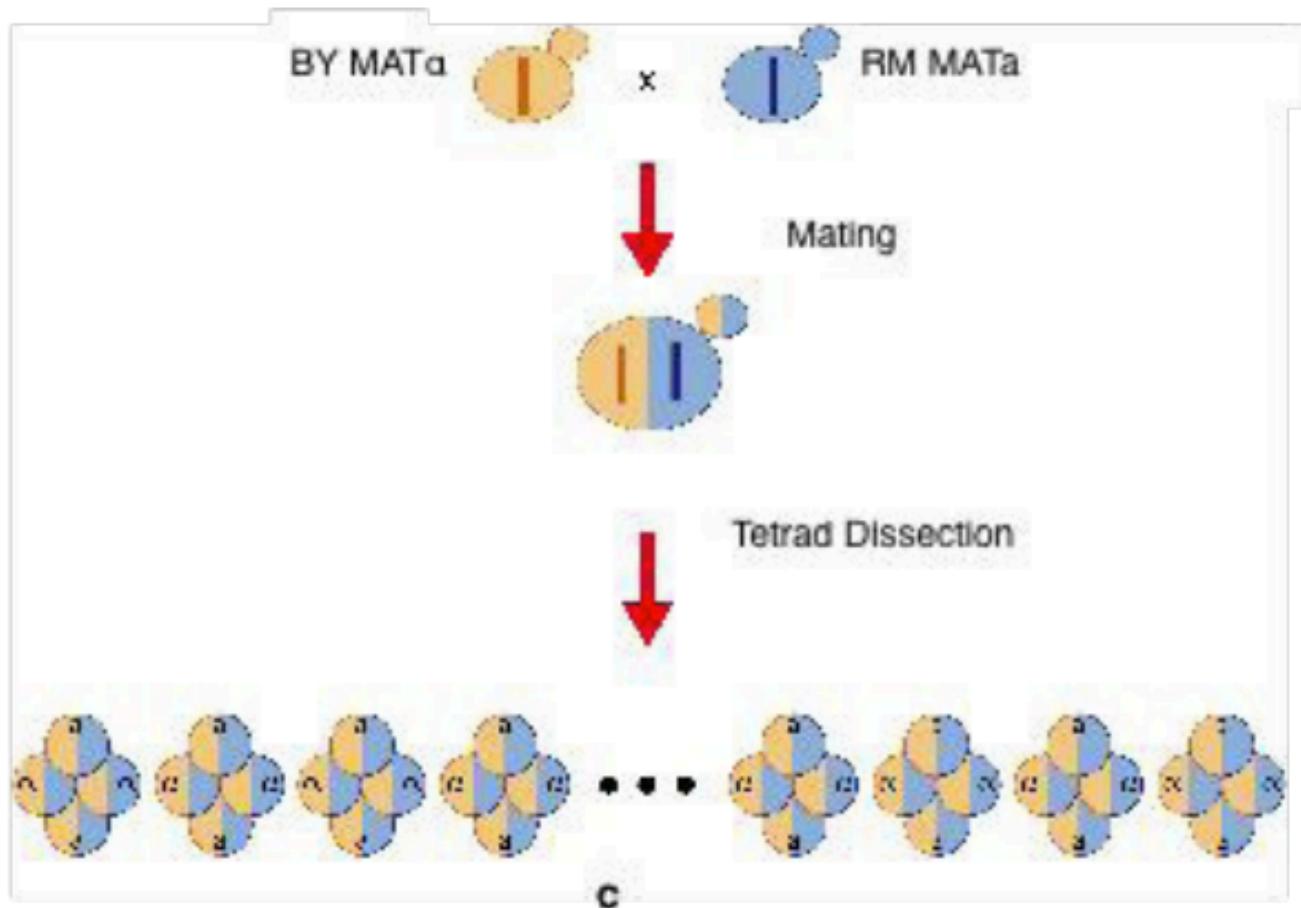
nature

# Methods

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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# 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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

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
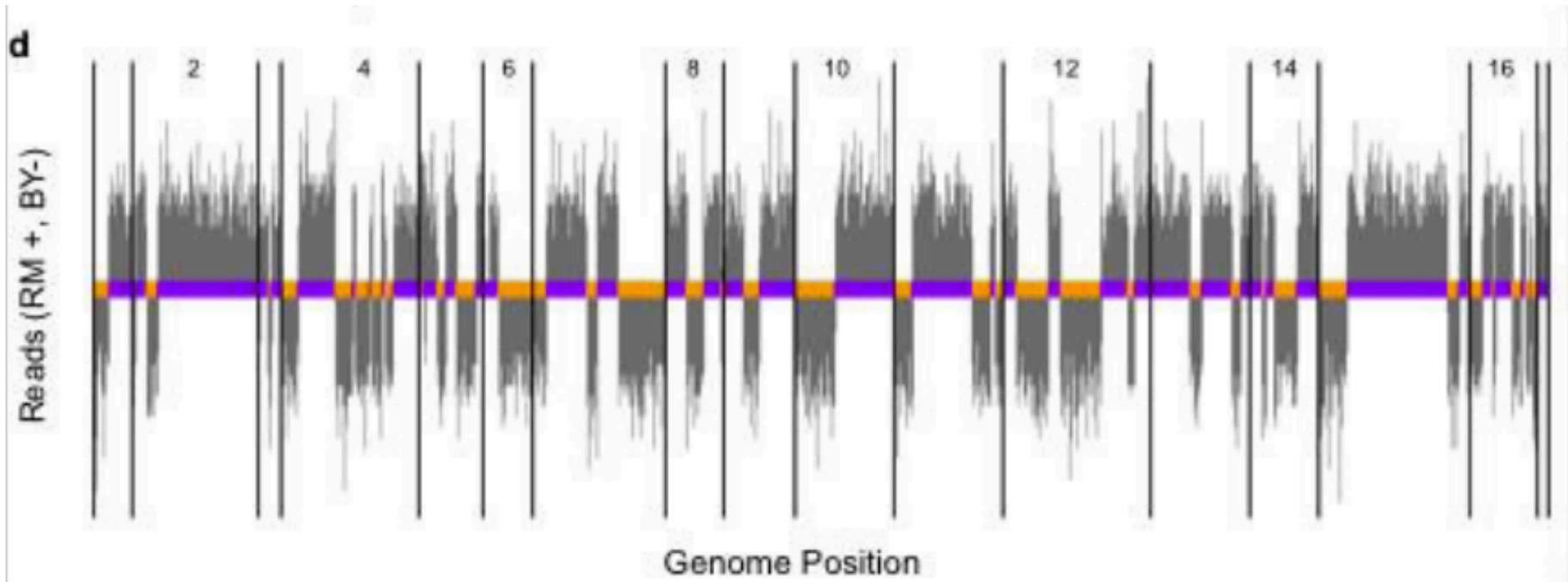
nature

# Methods

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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# 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: $\sqrt{\dfrac{pixelcount}{\pi}}$.

- QC with certain thresholds

- a robust locally weighted regression was fit to the radius measurements

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

(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.
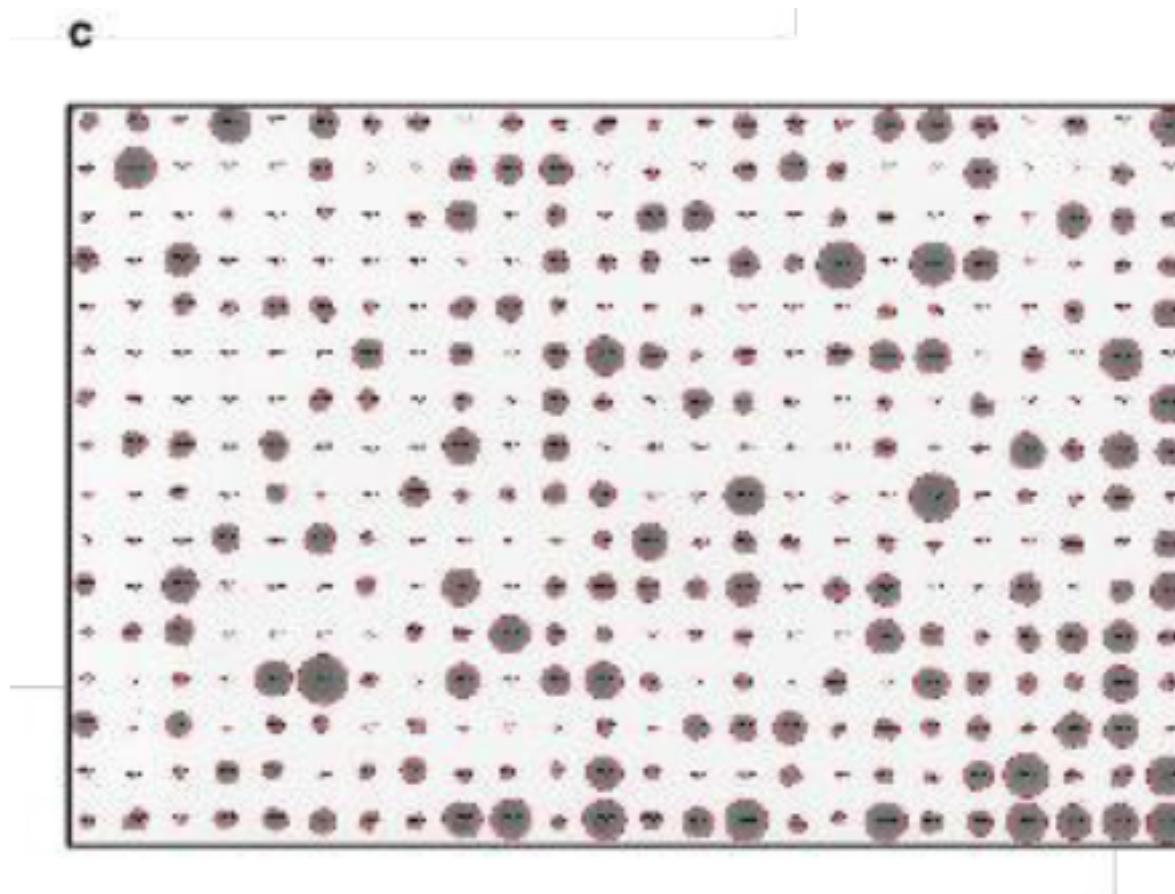


Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).
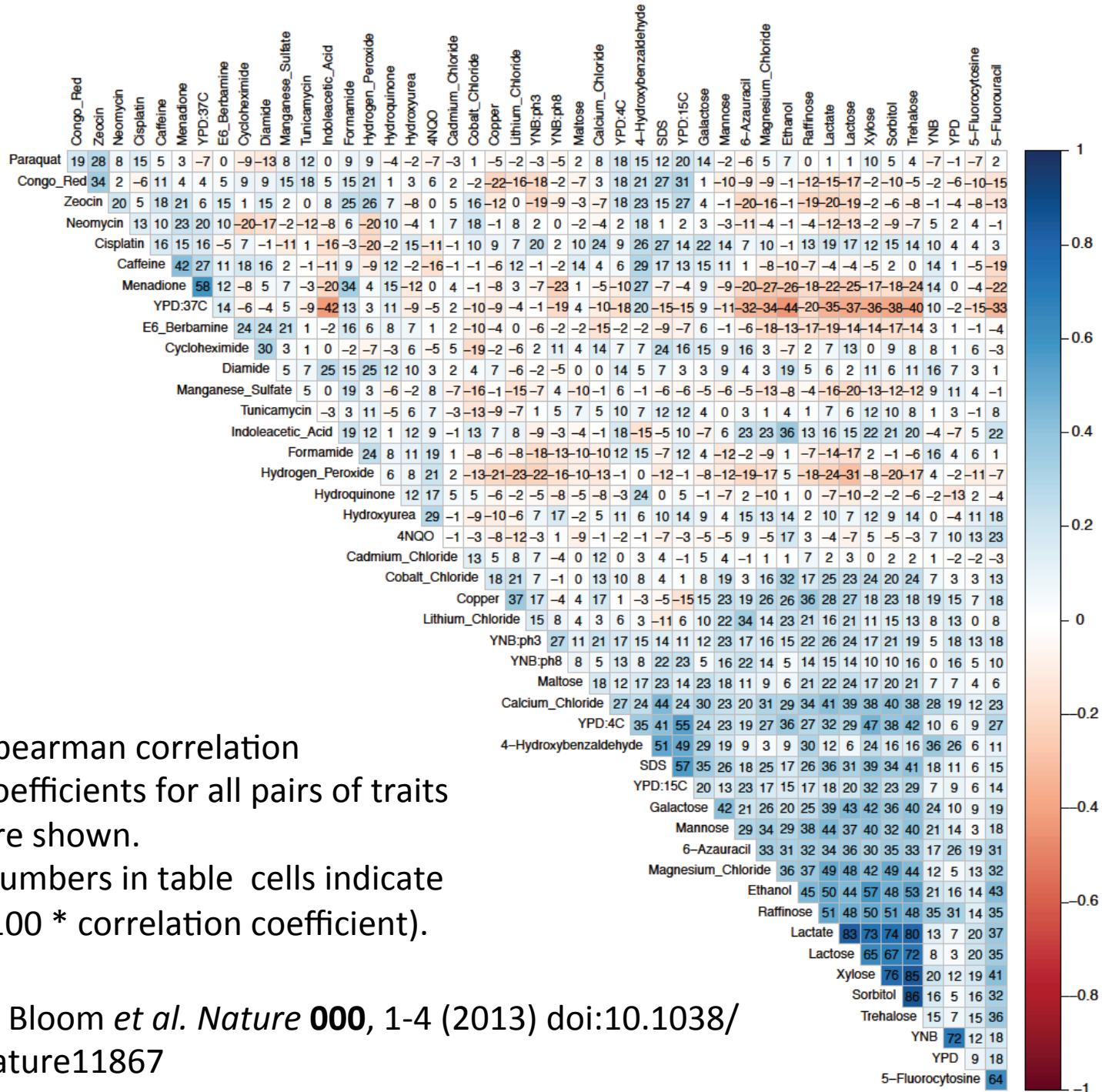
Spearman correlation
coefficients for all pairs of traits
are shown.
Numbers in table cells indicate
(100 * correlation coefficient).

nature

# Methods

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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# 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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# Methods

- Construction of segregant panel (1008)

- Sequencing step

- Phenotyping by end-point growth on agar plates

- Definition of genetic factors

- **Calculating heritability**

- QTL mapping

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# 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.

$$y = \beta 1_n + Zu + e \rightarrow V = A\sigma_A^2 + I\sigma_{EV}^2 \rightarrow \frac{\sigma_A^2}{\sigma_A^2 + \sigma_{EV}^2}$$

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

Heritability for 46 yeast traits.

nature

# Methods

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

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

# 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 minimumof 5, a maximumof 29 and a median of 12 QTL per trait

Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Nature 494, 234–237 (2013).

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.)

nature

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).
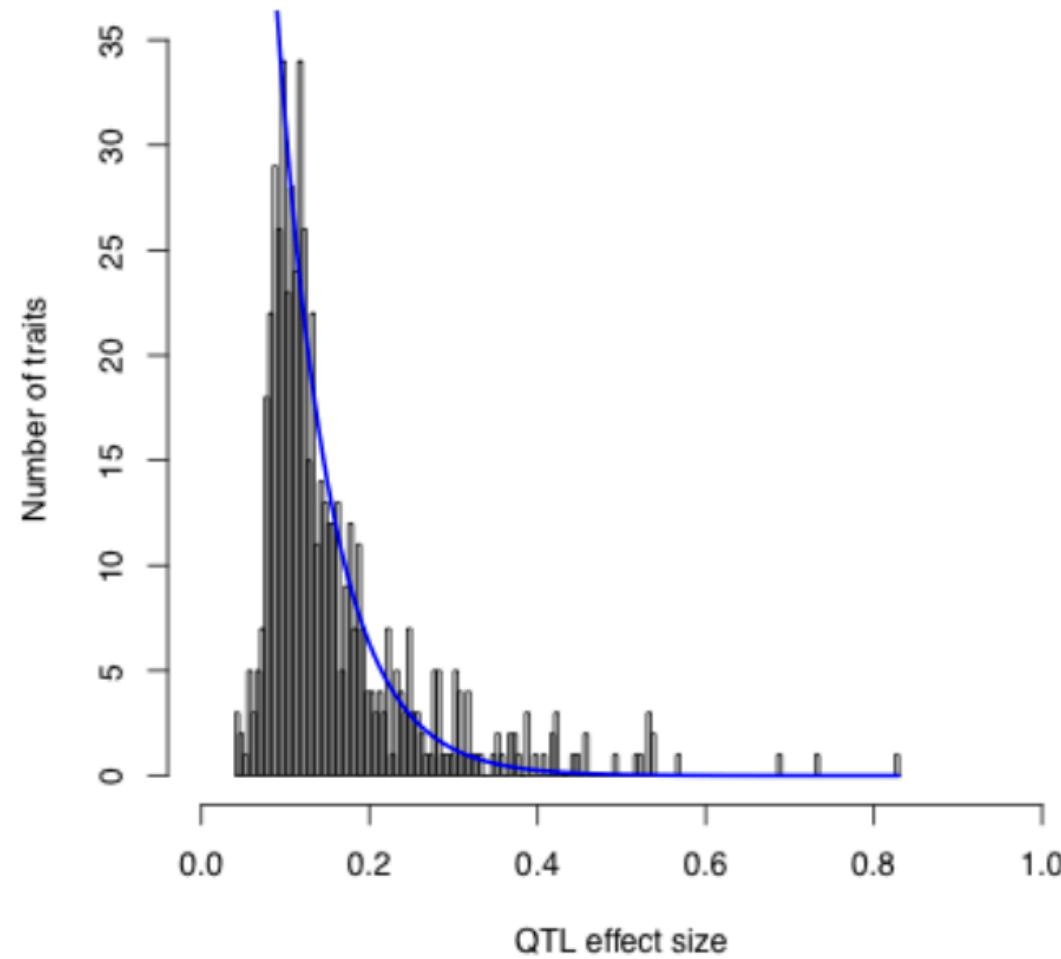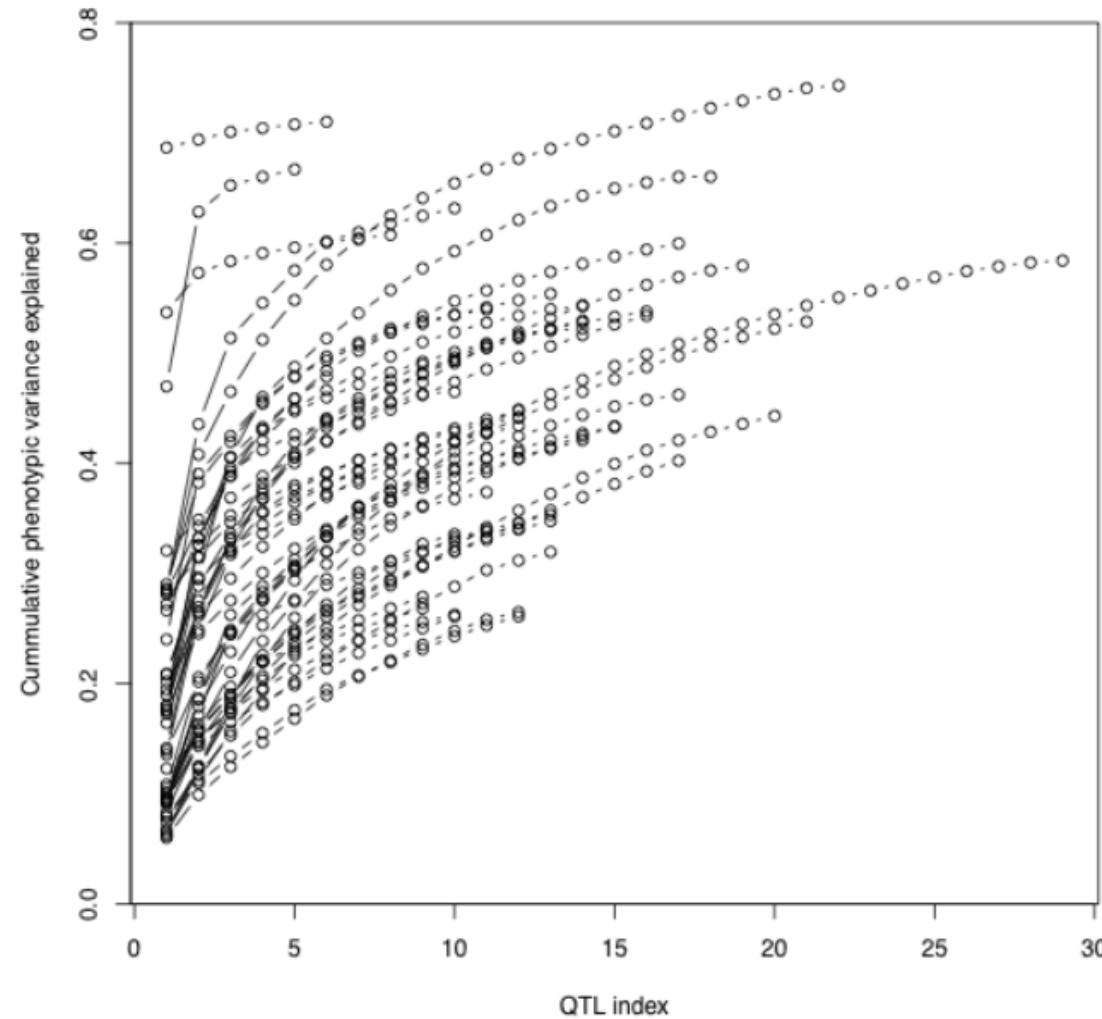


JS Bloom *et al.* Nature **000**, 1-4 (2013) doi:10.1038/
nature11867

nature

# Most additive heritability is explained by detected QTL.

nature

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.

nature

# Prediction of segregant trait values from QTL phenotypes.

nature

# 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 minimumof 1 and a maximum of 16 pairwise interactions per trait

# Non-additive genetic variance explained by QTL–QTL interactions.

nature

# 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

Frank W. Albert*, Sebastian Treusch, Arthur H. Shockley, Joshua S. Bloom and Leonid Kruglyak* (In press)

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

| Gene | X-pQTL LOD | eQTL LOD |
|---|---|---|
| Local eQTL only | | |
| YJL201W | 0.5 | 15.2 |
| YPL048W | 0.4 | 7.3 |
| YDL171C | 0.5 | 6.4 |
| YLR438W | 1.0 | 6.4 |
| YNL044W | 0.5 | 5.3 |
| Local X-pQTL only | | |
| YJL130C | 6.4 | 0.2 |
| YDL126C | 13.7 | 0.2 |
| YGL026C | 8.6 | 0.1 |
| YMR315W | 12.7 | 0.6 |

Albert et.al. In presss

**Table 2 – Hotspot regulators of protein expression**

| chromosome | Position (peak SNP) | % of genes regulated at LOD > 4.5 / LOD > 3 | mRNA hotspot[1] |
|---|---|---|---|
| I | 39,010 | 31 / 40 | Glu1 |
| II | 132,948 | 31 / 41 | - |
| II | 397,978 | 9 / 18 | Glu2 |
| IV | 223,943 | 12 / 24 | - |
| V | 192,064 | 16 / 31 | - |
| V | 371,845 | 16 / 21 | Glu6 |
| VII | 137,332 | 15 / 26 | - |
| VII | 505,871 | 16 / 29 | - |
| VIII | 103,041 | 19 / 29 | Glu7 |
| VIII | 419,747 | 8 / 12 | - |
| X | 142,009 | 18 / 26 | - |
| X | 655,465 | 11 / 15 | - |
| XI | 234,462 | 16 / 23 | Glu8 |
| XII | 238,302 | 16 / 31 | - |
| XII | 656,893 | 41 / 49 | Glu9 |
| XII | 1,039,502 | 12 / 19 | Yvert[2] |
| XIII | 96,832 | 31 / 46 | Glu10 |
| XIV | 232,509 | 13 / 19 | - |
| XIV | 465,007 | 58 / 65 | Glu11 |
| XV | 162,766 | 56 / 70 | Glu12 |

[1] As identified in Smith & Kruglyak 2008 [11].

[2] This hotspot was not observed in Smith & Kruglyak [11], but was present in an earlier BY / RM eQTL dataset[47].

Albert et.al. In presss

Table 3 – Genes regulated by the four hotspots discussed in the text

| Gene | chrXI effect | chrXII effect | Chr XV effect | Description |
|------|-------------|---------------|---------------|-------------|
| ATP14* | -0.35 | -0.14 | -0.14 | ATP synthase |
| ATP17* | -0.14 | -0.14 | -0.18 | ATP synthase |
| ATP2* | -0.21 | -0.3 | -0.22 | ATP synthase |
| CIT1* | -0.23 | -0.36 | -0.26 | Citrate synthase |
| MDH1* | -0.22 | -0.1 | -0.39 | Malate Dehydrogenase |
| ADO1 | -0.09 | -0.25 | 0.09 | Adenosine kinase |
| GLT1 | -0.08 | 0.13 | 0.24 | Glutamate synthase |
| LIA1 | -0.1 | 0.15 | 0.15 | Deoxyhypusine hydroxylase |
| TDH3 | -0.14 | 0.35 | 0.27 | Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) |
| YHB1 | -0.16 | -0.92 | 0.13 | Nitric oxide oxidoreductase |
| YLR179C | -0.09 | 0.7 | 0.17 | Unknown function |

* involved in aerobic respiration

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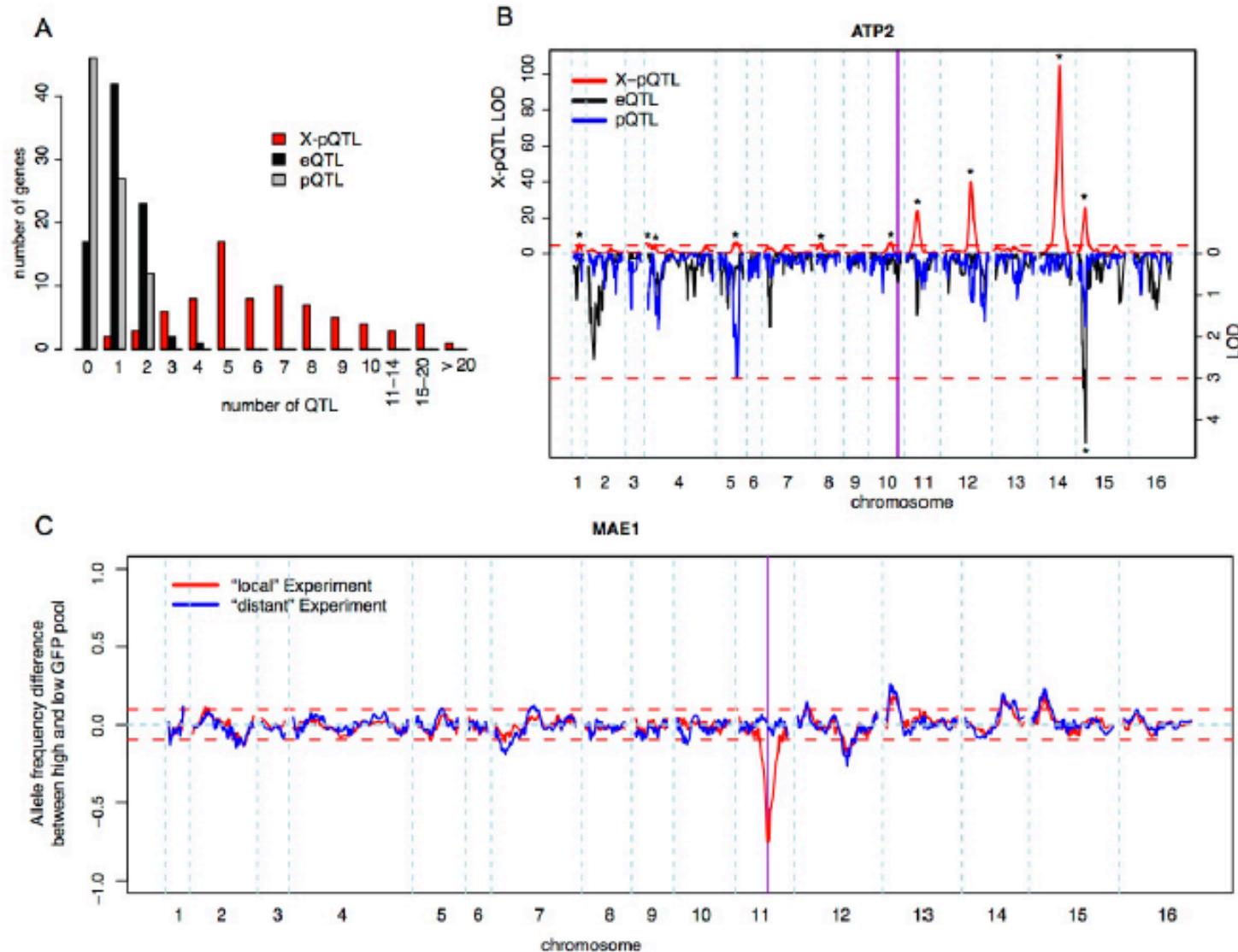
Figure 1 – Distant and local variation affects protein levels

Albert et.al. In presss
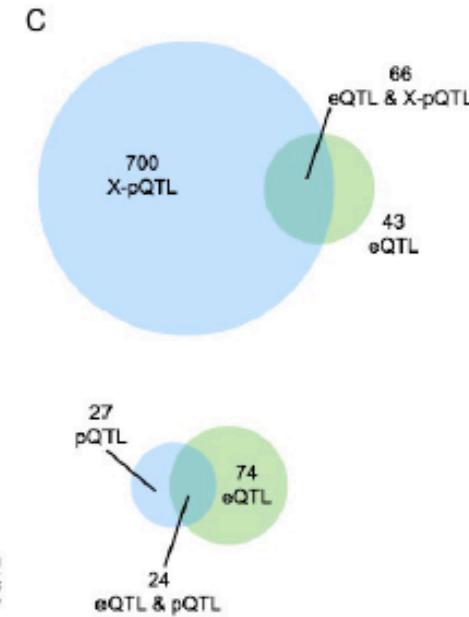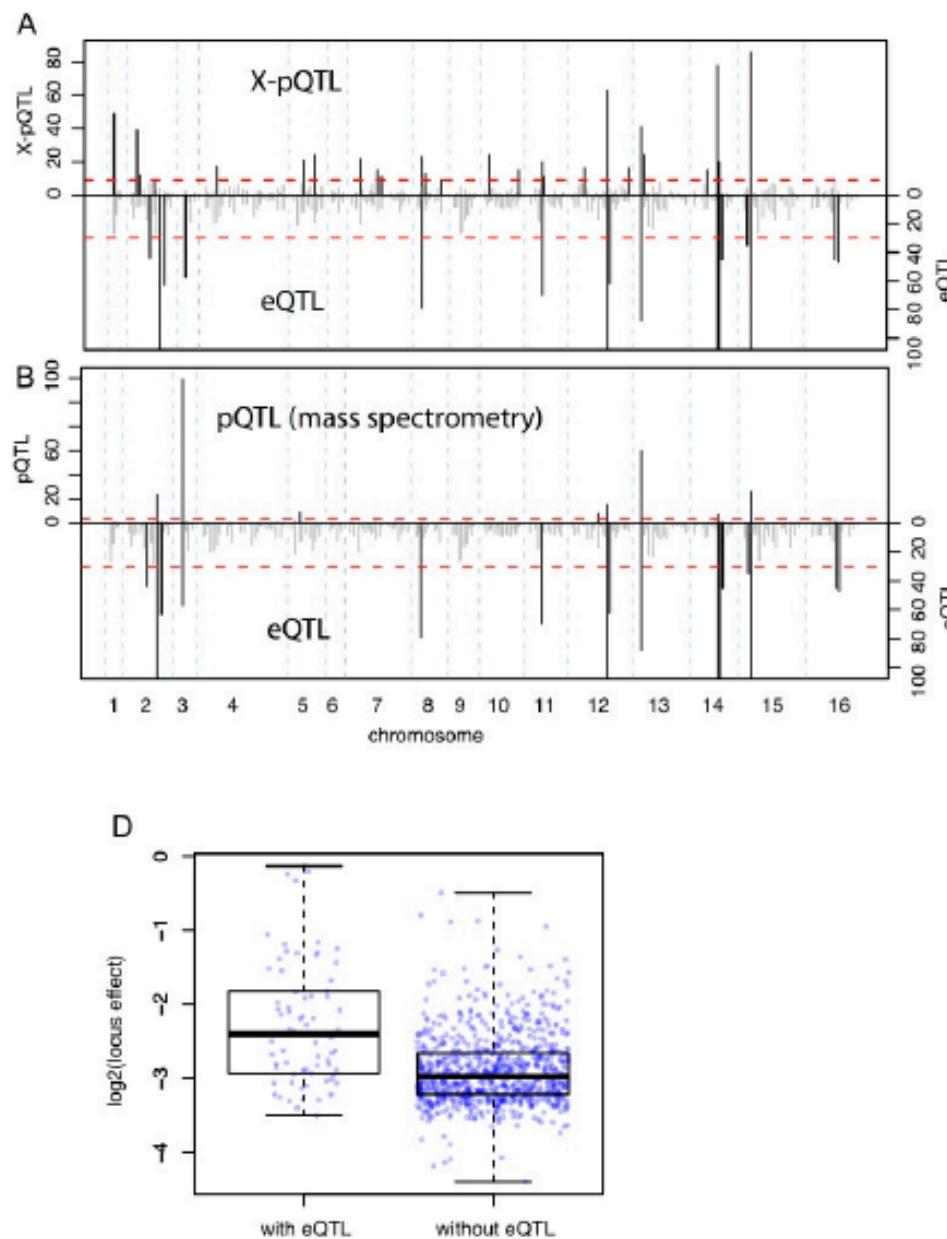
Figure 2 – X-pQTL hotspots and overlap with loci affecting mRNA abundance
Albert et.al. In presss
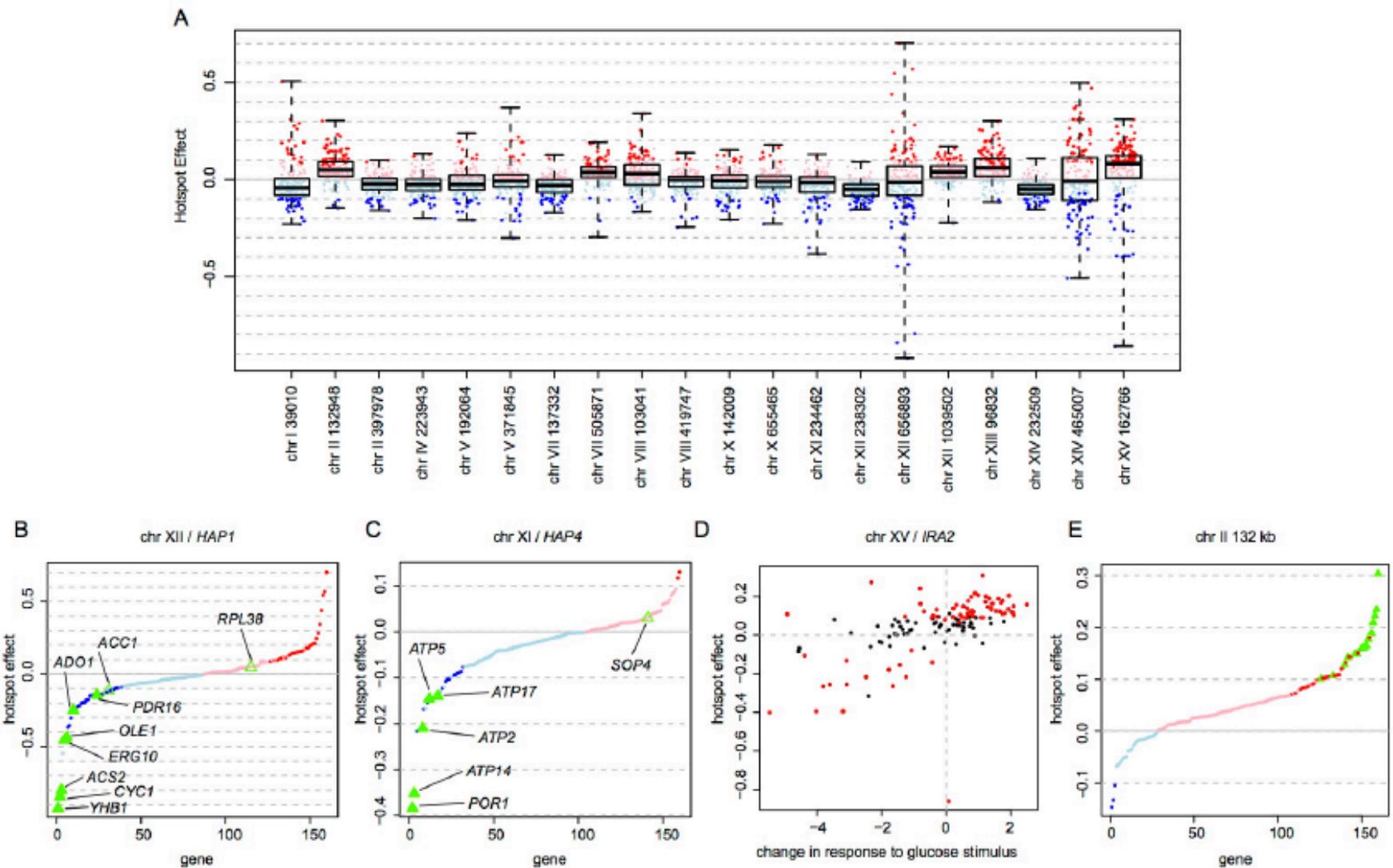
Figure 3 – Hotspot effects

Albert et.al. In presss
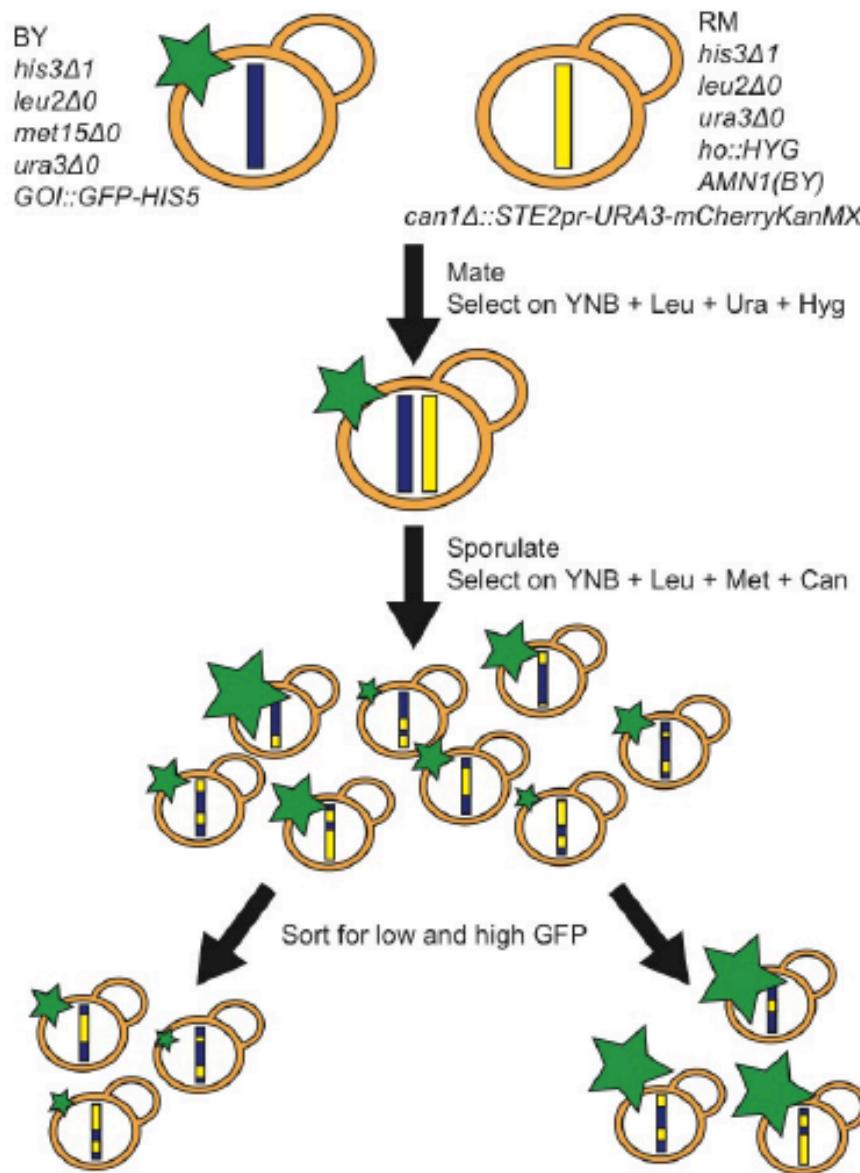
Fig. S1 Overview of the experimental design
Albert et.al. In presss
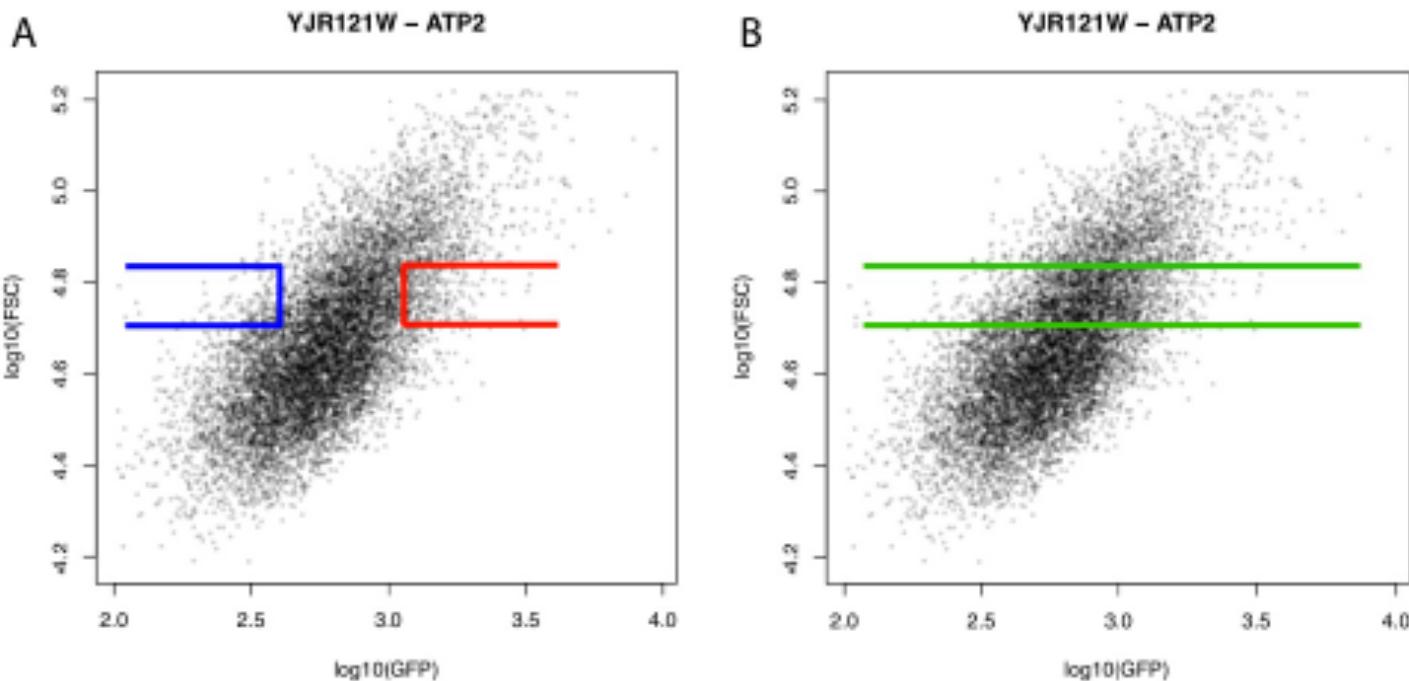
Supplementary Figure S2 – Illustration of FACS design

Albert et.al. In presss
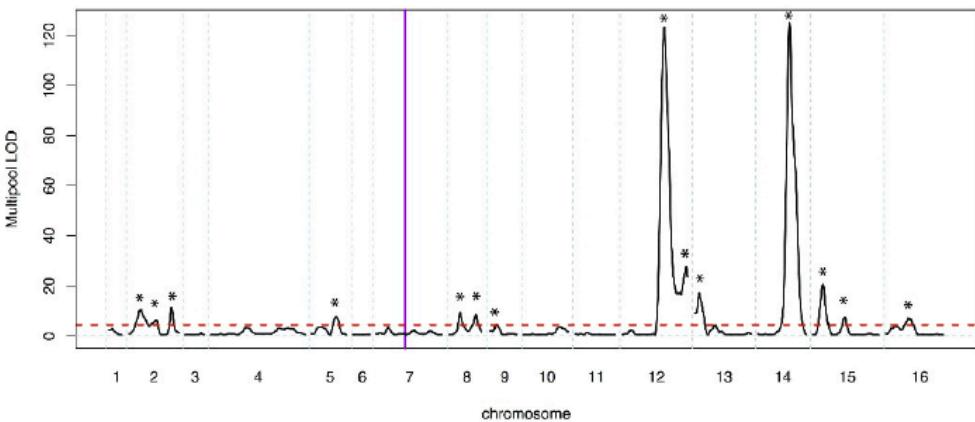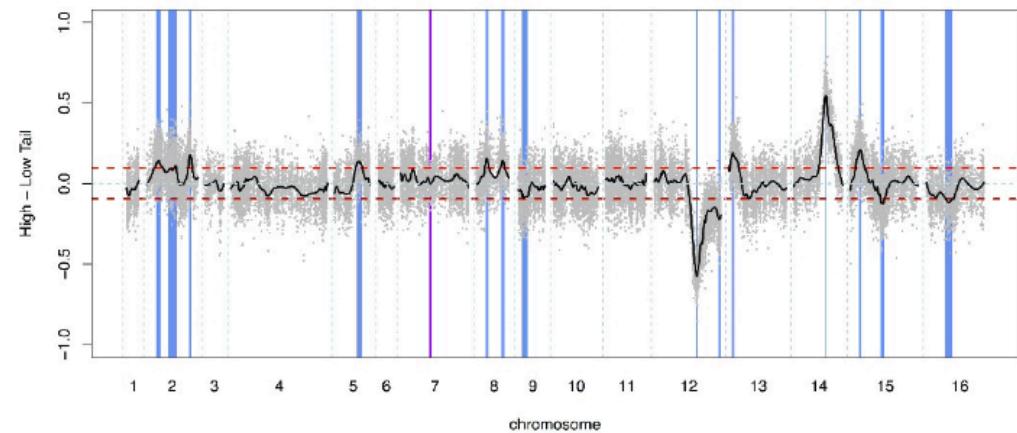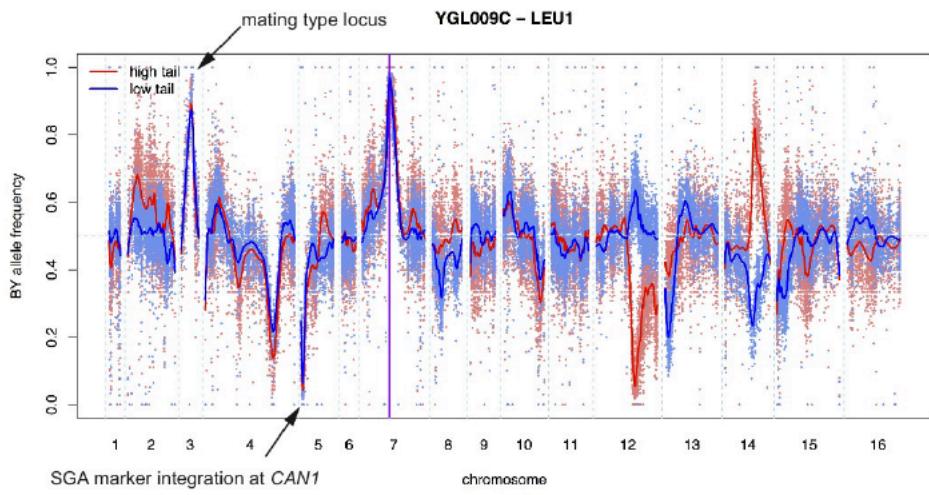
Supplementary Figure S3 –
Sequence analyses and X-pQTL
detection example

Albert et.al. In presss
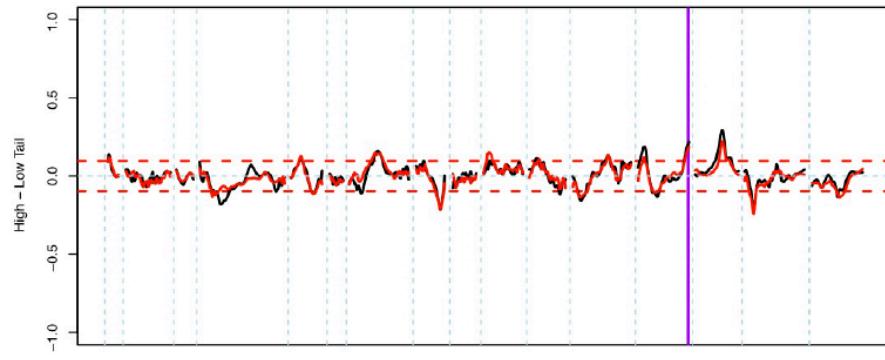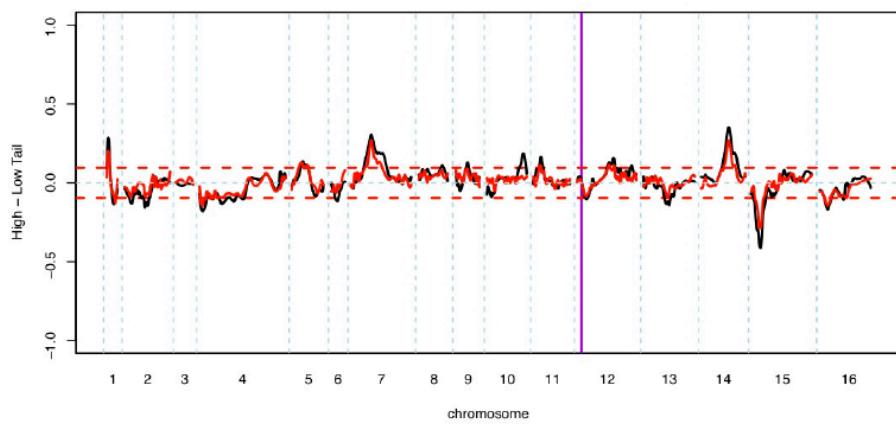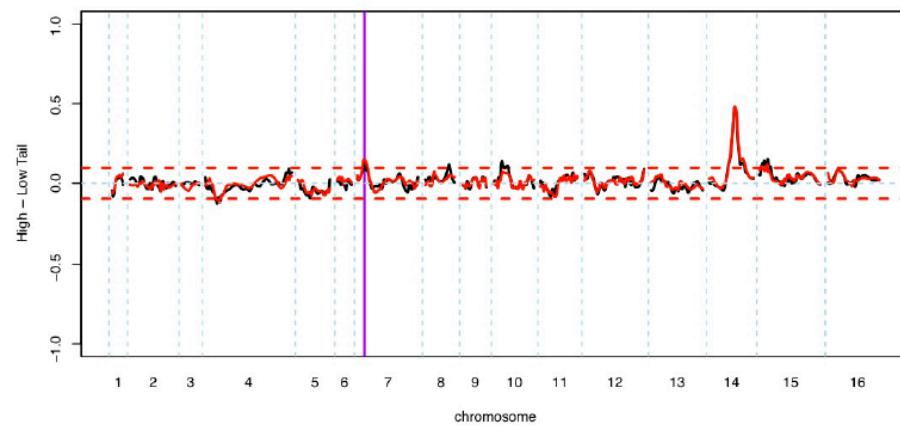
Supplementary Figure S4 – Reproducibility examples
Albert et.al. In presss
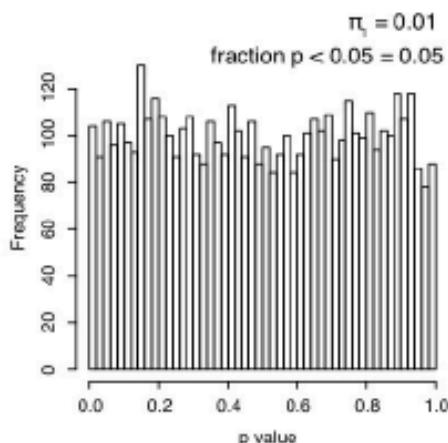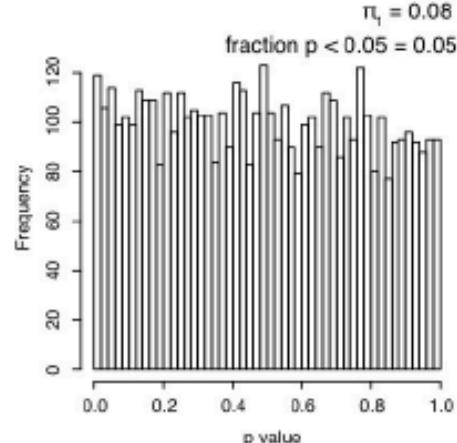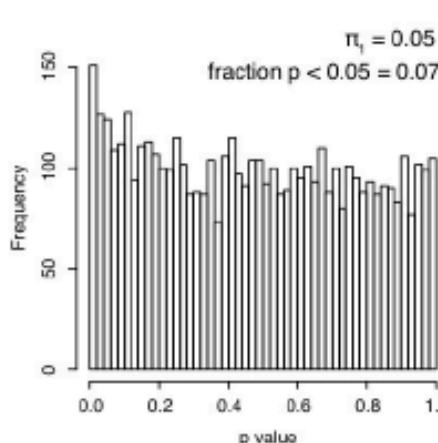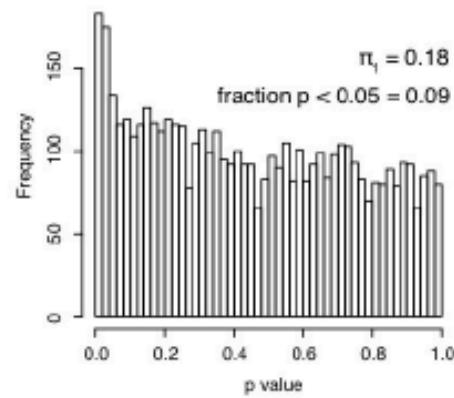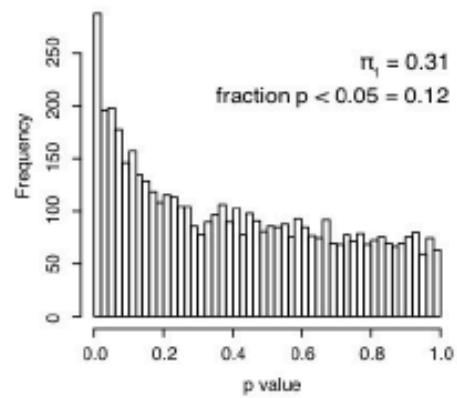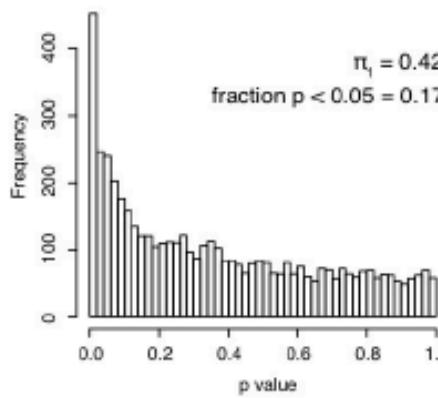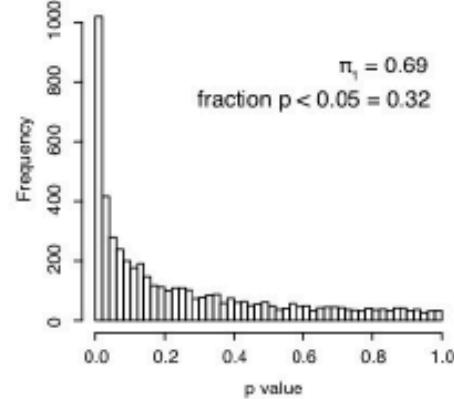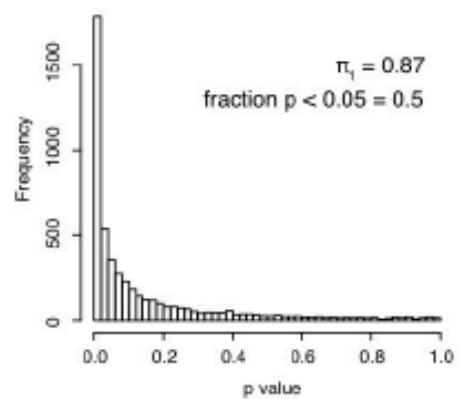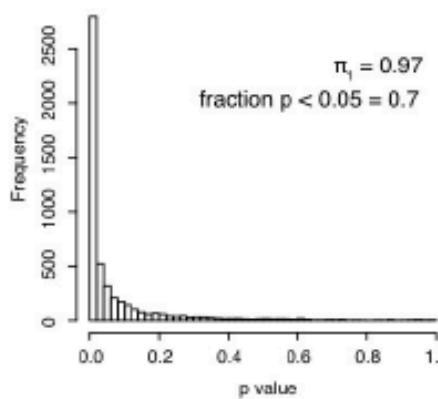
Supplementary Figure S5 – The impact of small effect sizes on the π1 estimate

Albert et.al. In presss
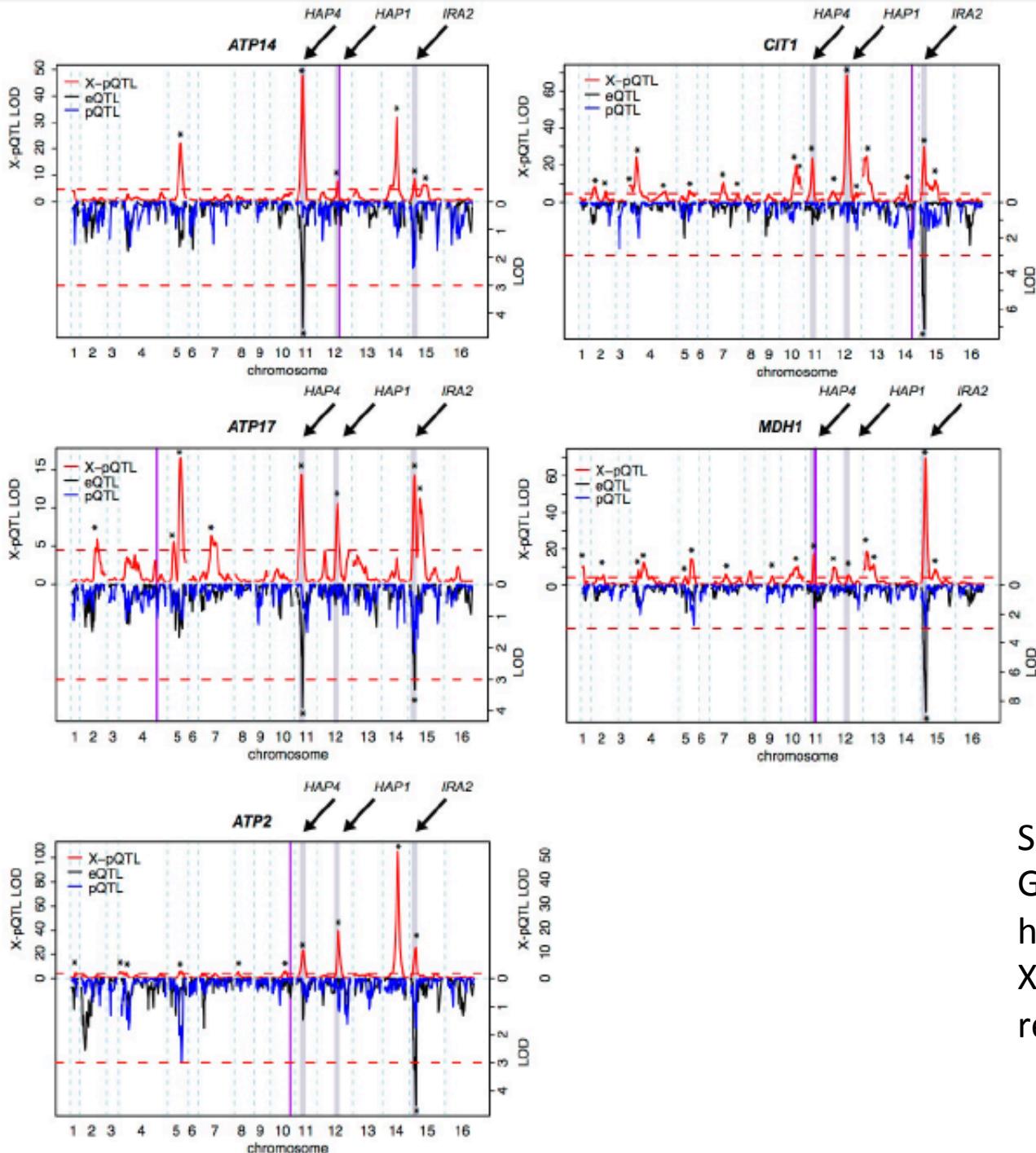
Supplementary Figure S6 – Genes regulated by the hotspots on chromosomes XI, XII, and XV involved in aerobic respiration

Albert et.al. In presss

# References

- Miles, C. & Wayne, M. (2008) Quantitative trait locus (QTL) analysis. Nature Education 1(1)

- Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. Nature 494, 234–237 (2013).

- Frank W. Albert*, Sebastian Treusch, Arthur H. Shockley, Joshua S. Bloom and Leonid Kruglyak*. Genetics of single-cell protein abundance variation in large yeast populations  (In press)