Supplementary information for

The Genetic Basis of Differential Autodiploidization in Evolving Yeast Populations

Sudipta Tung, Christopher W. Bakerlee, Angela M. Phillips, Alex N. Nguyen Ba, Michael M. Desai

Text S1. Data clean up prior to QTL analysis

Based on standard recommendations (Broman and Sen 2009), prior to QTL mapping following diagnostic probes were computed to ensure quality and integrity of the dataset.

Segregation distortion: Under normal circumstances BY and W303 alleles for each locus should segregate equally. To test this, we inspected genotype frequencies at each marker locus using function `geno.table`. 30 loci failed $\chi 2$ test for deviation from Mendelian proportions (i.e. 1:1, here). They were dropped from subsequent analysis.

Compare individuals' genotypes: In order to identify pairs of segregants with unusually similar genotypes across all loci, we compared genotypes for each pair of individuals using the `comparegeno` function. One pair of segregants had >99% similarity in genotype identity, was detected as an outlier (Grubb's test: Q = 5.81, $p = 0.0002$) and therefore removed from the subsequent analysis.

Counting crossovers: The number of crossover events observed for each segregant was computed using the `countXO` function. The number of crossovers was found to be unreasonably high for one segregant (Grubb's test: Q = 9.68, $p < 10^{-16}$), and this segregant was removed from further analysis.

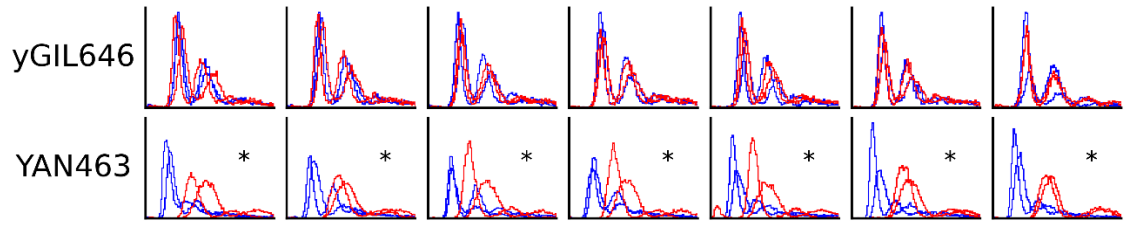

Figure S1. The ploidy state of the 7 replicate populations of the parental strains before and after 500-generation evolution. The plots show FITC histograms of Sytox-stained cells of each population, where the x-axis is in arbitrary fluorescence units (linear), and the y-axis is frequency. Blue and red curves denote the two technical replicate runs for each of the initial and final timepoints of evolution respectively. Populations where autodiploidization has been observed are marked by asterisks.

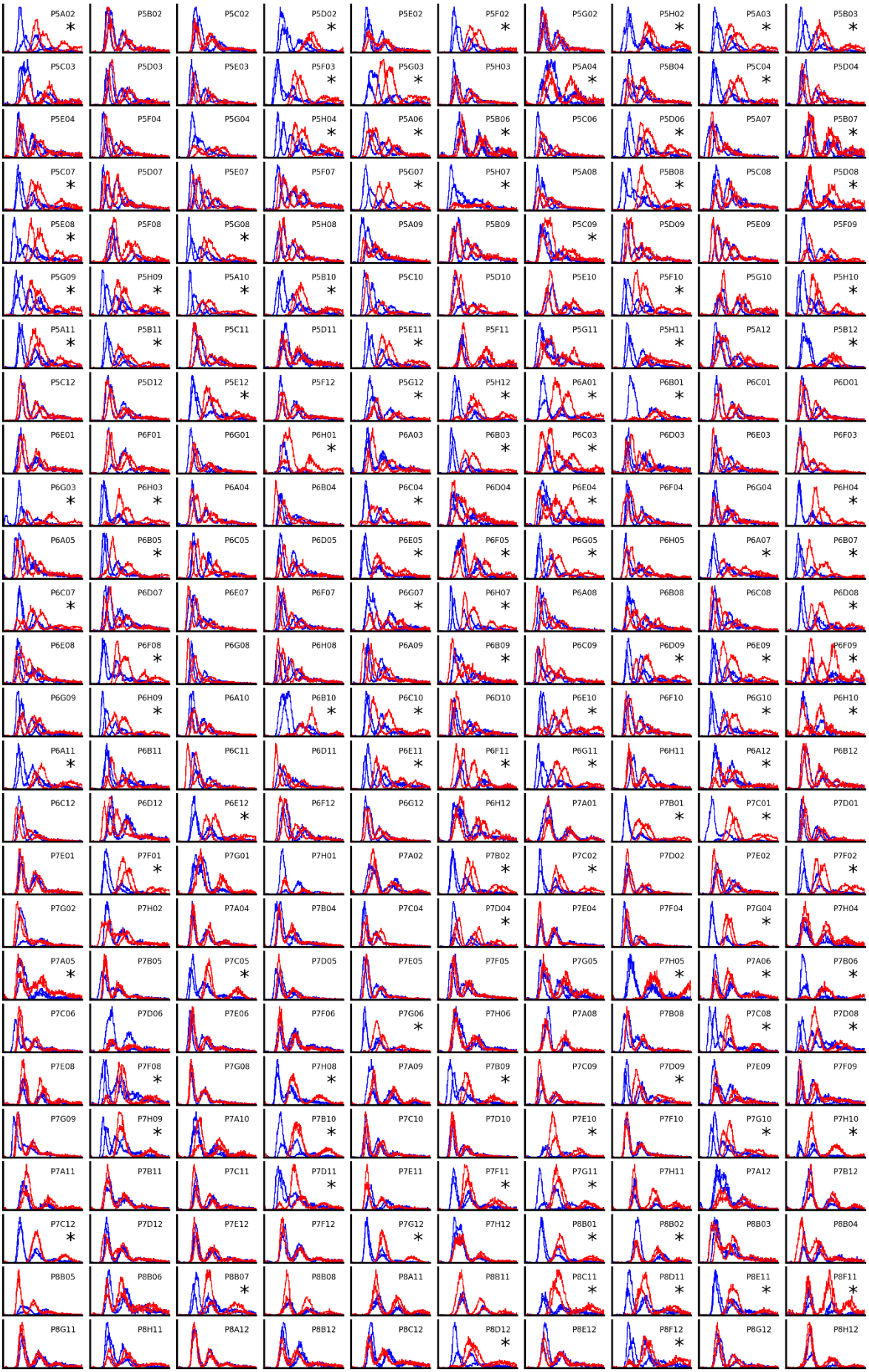
P5A02 *
P5B02
P5C02
P5D02 *
P5E02
P5F02 *
P5G02
P5H02 *
P5A03 *
P5B03 *
P5C03
P5D03
P5E03
P5F03 *
P5G03 *
P5H03
P5A04 *
P5B04
P5C04 *
P5D04
P5E04
P5F04
P5G04
P5H04 *
P5A06 *
P5B06 *
P5C06
P5D06 *
P5A07
P5B07 *
P5C07 *
P5D07
P5E07
P5F07
P5G07 *
P5H07 *
P5A08
P5B08 *
P5C08
P5D08 *
P5E08 *
P5F08
P5G08 *
P5H08
P5A09
P5B09
P5C09 *
P5D09
P5E09
P5F09
P5G09 *
P5H09 *
P5A10 *
P5B10 *
P5C10
P5D10
P5E10
P5F10 *
P5G10
P5H10 *
P5A11 *
P5B11 *
P5C11
P5D11
P5E11 *
P5F11
P5G11
P5H11 *
P5A12
P5B12 *
P5C12
P5D12
P5E12 *
P5F12
P5G12 *
P5H12 *
P6A01 *
P6B01 *
P6C01
P6D01
P6E01
P6F01
P6G01
P6H01 *
P6A03
P6B03 *
P6C03 *
P6D03
P6E03
P6F03
P6G03 *
P6H03 *
P6A04
P6B04
P6C04 *
P6D04
P6E04 *
P6F04
P6G04
P6H04 *
P6A05
P6B05 *
P6C05
P6D05
P6E05 *
P6F05 *
P6G05 *
P6H05
P6A07 *
P6B07 *
P6C07 *
P6D07
P6E07
P6F07
P6G07 *
P6H07 *
P6A08
P6B08
P6C08
P6D08 *
P6E08
P6F08 *
P6G08
P6H08
P6A09
P6B09 *
P6C09
P6D09 *
P6E09 *
P6F09 *
P6G09
P6H09 *
P6A10
P6B10 *
P6C10 *
P6D10
P6E10 *
P6F10
P6G10 *
P6H10 *
P6A11 *
P6B11
P6C11
P6D11
P6E11 *
P6F11 *
P6G11 *
P6H11
P6A12 *
P6B12
P6C12
P6D12
P6E12 *
P6F12
P6G12
P6H12
P7A01
P7B01 *
P7C01 *
P7D01
P7E01
P7F01 *
P7G01
P7H01
P7A02
P7B02 *
P7C02 *
P7D02
P7E02
P7F02 *
P7G02
P7H02
P7A04
P7B04
P7C04
P7D04 *
P7E04
P7F04
P7G04 *
P7H04
P7A05 *
P7B05
P7C05 *
P7D05
P7E05
P7F05
P7G05
P7H05 *
P7A06 *
P7B06 *
P7C06
P7D06
P7E06
P7F06
P7G06 *
P7H06
P7A08
P7B08
P7C08 *
P7D08 *
P7E08
P7F08 *
P7G08
P7H08 *
P7A09
P7B09 *
P7C09
P7D09 *
P7E09
P7F09
P7G09
P7H09 *
P7A10
P7B10 *
P7C10
P7D10
P7E10 *
P7F10
P7G10 *
P7H10 *
P7A11
P7B11
P7C11
P7D11 *
P7E11
P7F11 *
P7G11 *
P7H11
P7A12
P7B12
P7C12 *
P7D12
P7E12
P7F12
P7G12 *
P7H12
P8B01 *
P8B02 *
P8B03
P8B04
P8B05
P8B06
P8B07 *
P8B08
P8A11
P8B11
P8C11 *
P8D11 *
P8E11 *
P8F11 *
P8G11
P8H11
P8A12
P8B12
P8C12
P8D12 *
P8E12
P8F12 *
P8G12
P8H12

Figure S2. The ploidy state of the 260 tetrad populations before and after 500-generation evolution. The plots show FITC histograms of Sytox-stained cells of each population, where the x-axis is in arbitrary fluorescence units (linear), and the y-axis is frequency. Code starting with 'P' on each panel indicates population ID. Blue and red curves denote the two technical replicate runs for each of the initial and final timepoints of evolution respectively. Populations where autodiploidization has been observed are marked by asterisks.

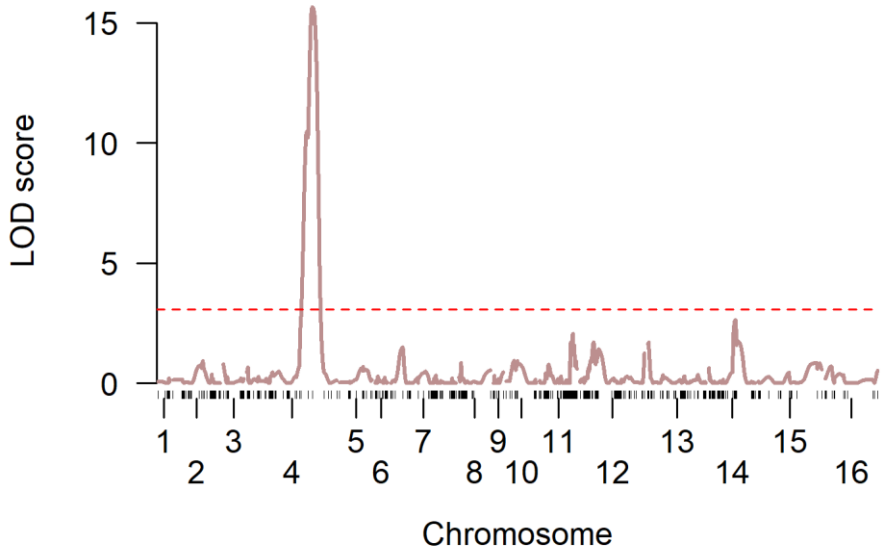

Figure S3. LOD score for variation in autodiploidization obtained using the Haley–Knott regression method is plotted against the genetic map. The red dashed line indicates a 5% LOD significance threshold computed from 10,000 permutations. The single statistically significant QTL is identical to that of Figure 2C and falls within the *SSD1* locus.

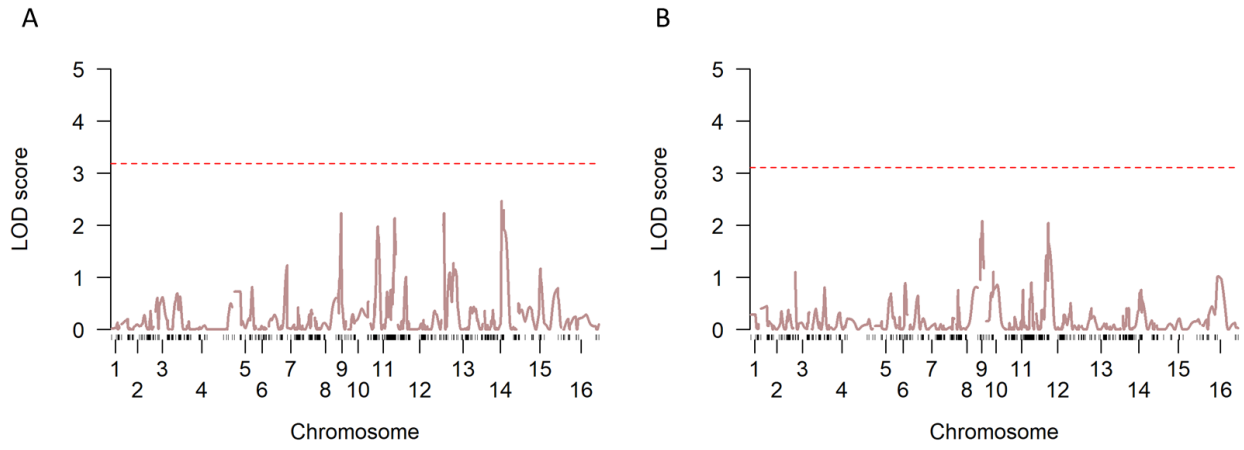

Figure S4. LOD score for variation in autodiploidization obtained using standard interval mapping method after regressing out the statistically significant chromosome IV QTL. The red dashed line indicates a 5% LOD significance threshold computed from 10,000 permutations. No additional statistically significant QTL are present for the segregants with (A) BY and (B) W303 allele of the chromosome IV QTL.

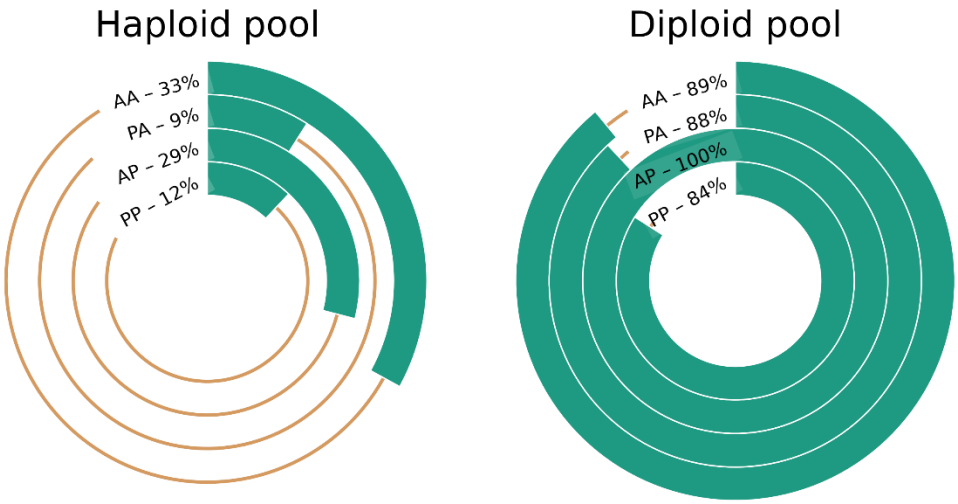

Figure S5. Percentage of sequencing reads at *SSD1* locus matching BY allele in haploid and diploid pools of the 'selected spores.' The two-letter code for each plot indicate whether they are auxotrophic (A) or prototrophic (P) for Tryptophan and Lysine, (e.g. 'PA' denotes the spores that are prototrophic for Tryptophan but auxotrophic for Lysine). Irrespective of the auxotrophy status, the BY allele is substantially enriched in the diploid pool, whereas it is depleted in the haploid pool.

Table S1. Occurrence of autodiploidization in "tetrad spores," categorized by prototrophy or auxotrophy for tryptophan, uracil, and lysine.

Tryptophan				
	Num. haploid	% haploid	Num. diploid	% diploid
Auxotrophic	80	63%	47	37%
Prototrophic	65	50%	66	50%
Uracil				
	Num. haploid	% haploid	Num. diploid	% diploid
Auxotrophic	69	54%	58	46%
Prototrophic	76	58%	55	42%
Lysine				
	Num. haploid	**% haploid**	**Num. diploid**	**% diploid**
Auxotrophic	62	49%	64	51%
Prototrophic	83	63%	49	37%

Table S2. Non-synonymous differences between SSD1 alleles of strains examined in this study.

Strain	Amino acid position in *SSD1* *total length = 1250 AAs*					
	377	693	698	1190	1196	1250
BY4741	S	T	Y	S	A	V
W303	S	T	*	S	A	V
RM11-1a	S	T	Y	S	A	V
DBVPG1106	S	T	Y	G	P	A
Y55	C	M	Y	G	P	V
SK1	C	M	Y	G	P	V
YPS128	S	T	Y	G	P	A

References

Broman, K.W., and S. Sen, 2009 *A Guide to QTL Mapping with R/qtl* New York: Springer.