Supplementary Material

Recombining your way out of trouble: The genetic architecture of hybrid fitness under environmental stress

Zebin Zhang[1*], Devin P. Bendixsen[1*], Thijs Janzen[2,3], Arne W. Nolte[2,3], Duncan Greig[2,4], Rike Stelkens[1,2]
*These authors contributed equally to the study.

[1]Division of Population Genetics, Department of Zoology, Stockholm University, Svante Arrheniusväg 18 B, 106 91 Stockholm, Sweden

[2]Max Planck Institute for Evolutionary Biology, August-Thienemann Straße 2, 24306 Plön, Germany

[3]Institute of Biology and Environmental Sciences, University of Oldenburg, Carl-von-Ossietzky-Straße 9-11, 26111 Oldenburg, Germany

[4]Centre for Life's Origins and Evolution (CLOE), Department of Genetics, Evolution, and Environment, University College London, Gower Street, London WC1E 6BT, United Kingdom

Table S1 Toxins and concentrations

Concentrations (in M or %) of the 10 toxins used in the experiment. Numbers 1-8 indicate row numbers from top (lethal) to bottom (permissive) on 96-well culture plates.

Substance		Concentration							
		1	2	3	4	5	6	7	8
salicylic acid ($C_7H_6O_3$)	M	0.086	0.057	0.038	0.025	0.017	0.011	0.008	0.005
caffeine ($C_8H_{10}N_4O_2$)	M	0.103	0.069	0.046	0.031	0.020	0.014	0.009	0.006
zinc sulfate ($ZnSO_4$),	M	0.186	0.124	0.083	0.055	0.037	0.024	0.016	0.011
Nipagin ($CH_3(C_6H_4(OH)COO)$)	M	0.053	0.035	0.024	0.016	0.010	0.007	0.005	0.003
sodium chloride (NaCl)	M	4.0	2.667	1.778	1.185	0.790	0.527	0.351	0.234
citric acid ($C_6H_8O_7$)	M	0.115	0.077	0.051	0.034	0.023	0.015	0.010	0.007
lithium acetate (CH_3COOLi)	M	0.5	0.333	0.222	0.148	0.099	0.066	0.044	0.029
hydrogen peroxide (H_2O_2)	%	0.014	0.009	0.006	0.004	0.003	0.002	0.001	0.001
ethanol (C_2H_6O)	%	50	33.333	22.222	14.815	9.877	6.584	4.390	2.926
DMSO (C_2H_6OS)	%	30	20.000	13.333	8.889	5.926	3.951	2.634	1.756

Table S2 Linear mixed-effect models used to predict zygosity

The most appropriate model was selected by identifying the simplest model that maintained the lowest Akaike Information Criteron (AIC). The optimal model, given our data, is indicated by a black square.

model	variable count	Response Variable	fixed effects			random effects	AIC	BIC
			Environment	Chromosome ID	Chromosome ID : Environment			
1	3	Zygosity	+	+	+	1|S	1899.3	2869.3
2	2		+	+		1|S	1921.3	2083
3			+		+	1|S	1899.3	2869.3
4				+	+	1|S	1899.3	2869.3
5	1		+			1|S	2659.9	2731.7
6				+		1|S	1914.6	2022.4
7					+	1|S	1899.3	2869.3

Table S3 Linear mixed-effect models used to predict chromosome hybridity

The most appropriate model was selected by identifying the simplest model that maintained the lowest Akaike Information Criteron (AIC). The optimal model, given our data, is indicated by a black square.

model	variable count	Response Variable	fixed effects										random effects	AIC	BIC
			A	E	C	P	A:C	A:E	A:P	C:E	C:P	A:C:P			
1	10	Hybridity	+	+	+	+	+	+	+	+	+	+	1|S	-1512.8	-177.3
2	9		+	+	+	+	+	+	+	+	+		1|S	-1510.7	-237.6
3			+	+	+	+	+	+	+	+		+	1|S	-1512.8	-177.3
4			+	+	+	+	+	+	+			+	1|S	-1399	-906
5	8		+	+	+	+	+	+		+		+	1|S	-1512.8	-177.3
6	7		+	+	+	+	+			+		+	1|S	-1506.5	-227.2
7			+	+	+	+		+		+		+	1|S	-1512.8	-177.3
8	6		+	+	+			+		+		+	1|S	-1512.8	-177.3
9	5		+	+				+		+		+	1|S	-1512.8	-177.3
10	4		+					+		+		+	1|S	-1512.8	-177.3
11	3							+		+		+	1|S	-1512.8	-177.3
12	2							+		+			1|S	-737.1	336.3
13								+				+	1|S	-1399	-906
14										+		+	1|S	-1506.5	-227.2
15	1											+	1|S	-1404.4	-1023.7
16	1									+			1|S	-727.2	283.8
17	1							+					1|S	1062.3	1199.6

A=Aneuploidy E=Environment C=Chromosome ID P=Ploidy

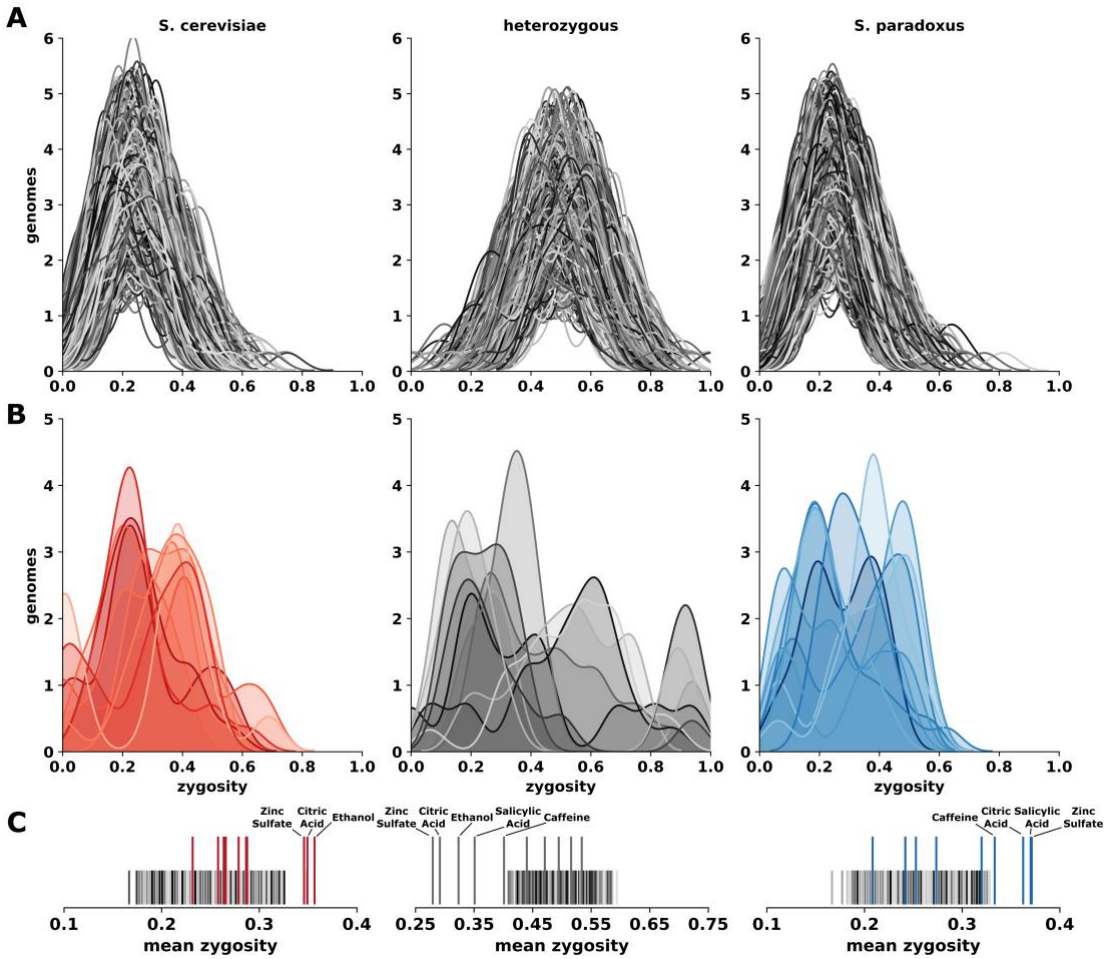

Figure S1 Genome-wide zygosity of simulated chromosomes
(A) The distributions of homozygosity for *S. cerevisiae* or *S. paradoxus* and heterozygosity in a simulated environment resulting from simulated chromosomes based on no chromosomal interactions (no epistasis) and random segregation. Each environment consists of 24 genomes, each consisting of 16 simulated chromosomes. Each simulated environment was repeated 10,000 times. **(B)** The distribution of zygosities found in our experimental environments. Distributions are colored according to their zygosity (red = homozygous *S. cerevisiae*, grey = heterozygous, blue = homozygous *S. paradoxus*). **(C)** Mean zygosity of our experimental environments as compared to the range of mean zygosities found in simulations (grey-black lines).

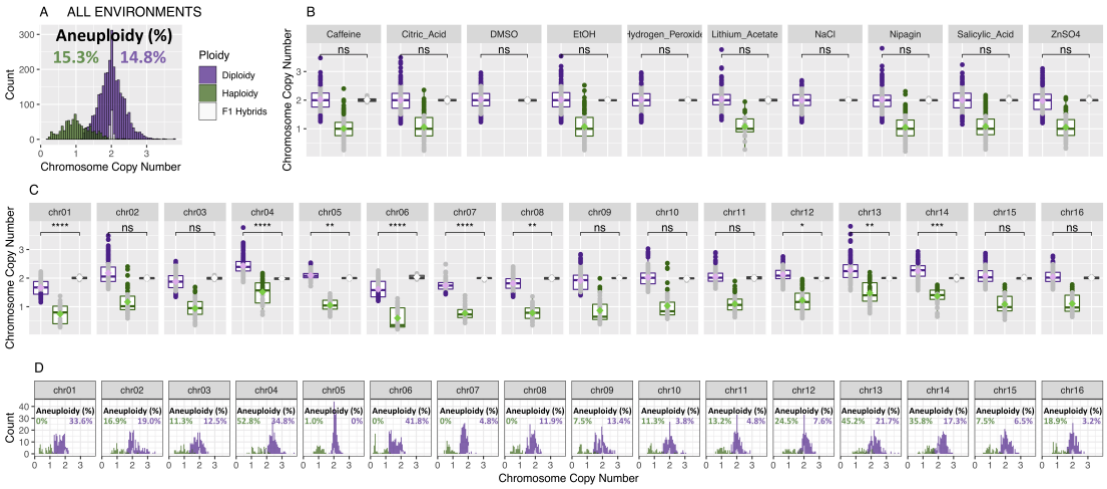

Figure S2 Chromosome karyotyping of hybrids

(A) Distribution of chromosome copy numbers found in 237 F2 hybrids. Overall diploid genomes are purple, overall haploid genomes are green, euploid F1 genomes used for normalization are white. Overall aneuploidy was 14.8% in diploids and 15.3% in haploids (chromosomes were considered aneuploid if the value was >2.5 and <1.5 in diploids, and >1.5 in haploids). **(B)** Chromosome copy numbers within environments. Lines in boxes are medians and diamonds indicate means. Each dot represents a chromosome. Grey indicates expected chromosome copy number, purple indicates aneuploidy in diploids, and green indicates aneuploidy in haploids. Wilcoxon tests comparing overall chromosome copy number of F2 hybrids and F1 hybrids were non-significant. **(C)** Chromosome copy numbers per chromosome. Asterisks indicate significant differences between F2 hybrids and F1 hybrids using Wilcoxon signed-rank tests (p = *0.05, **0.01, ***0.001, ****0.0001). **(D)** Chromosome copy number distribution shows large variation between chromosomes.

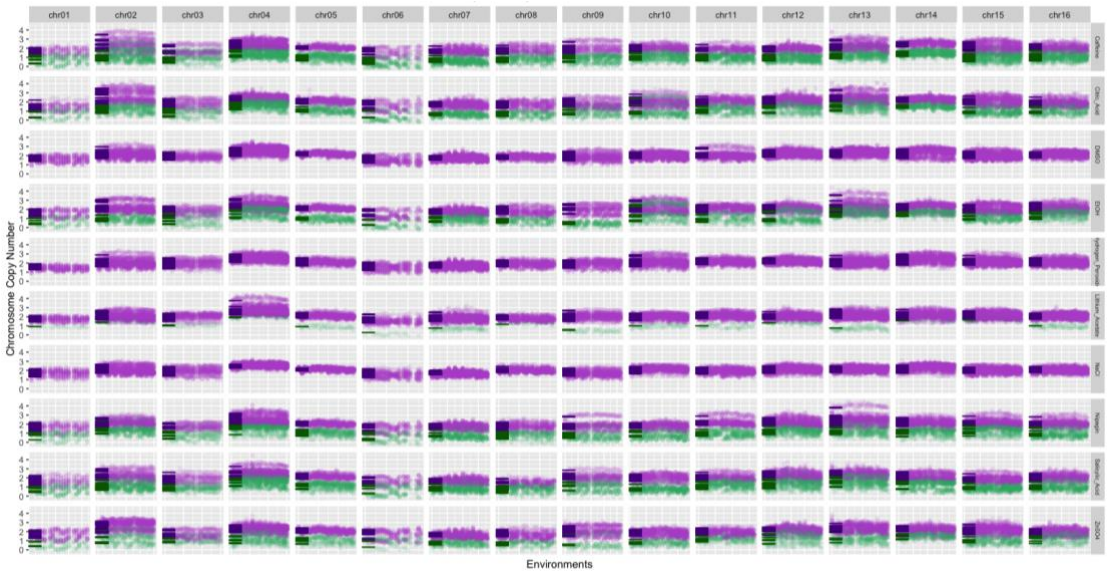

Figure S3 Landscape of chromosome copy number of 237 F2 hybrid genomes. Rows each show one of 10 stressful environments, columns each show one of 16 chromosomes. Each dot represents a 10kb bin along the chromosomes. Mean values of each chromosome are shown on the left side of each grid slightly darkened for emphasis. Purple indicates overall diploid genomes, green indicates overall haploid genomes.

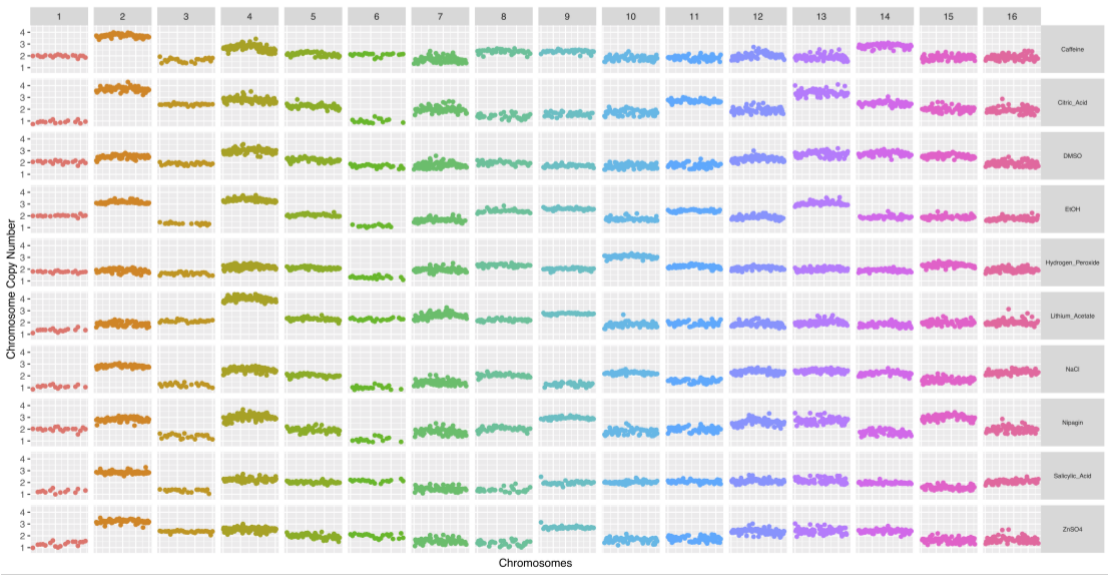

Figure S4 Example genomes from 10 stressful environments.
Each row shows chromosome copy numbers of one representative, aneuploid F2 hybrid genome chosen from each environment. Each dot represents a 10kb bin within each chromosome.

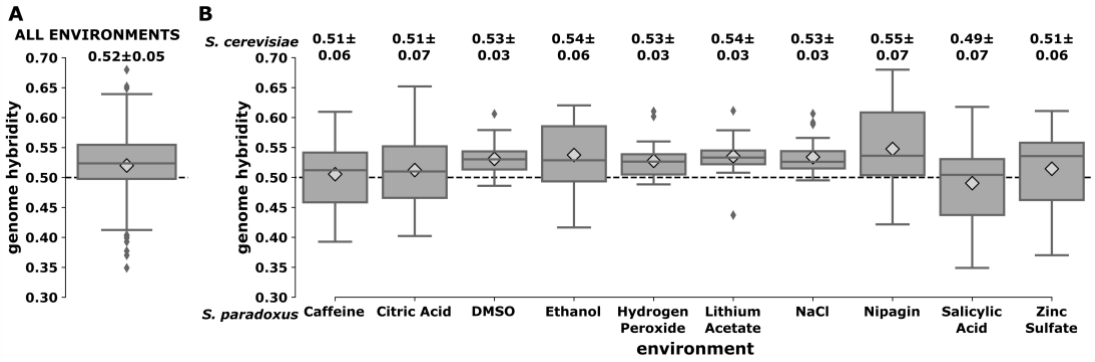

Figure S5 Genome-wide hybridity of 237 F2 hybrids
(A) Genome hybridity across all chromosomes and all environments, measured as the mean of the 16 chromosome hybridities (Figure 2) within a genome. Boxplots indicate the median and interquartile range. Diamonds indicate the mean. Dashed line (0.5) indicates equal amount of the genome mapping to *S. cerevisiae*. **(B)** Genome hybridity within environments. Numbers indicate the average genome hybridity for each F2 genome within each environment (n genomes ~ 24).

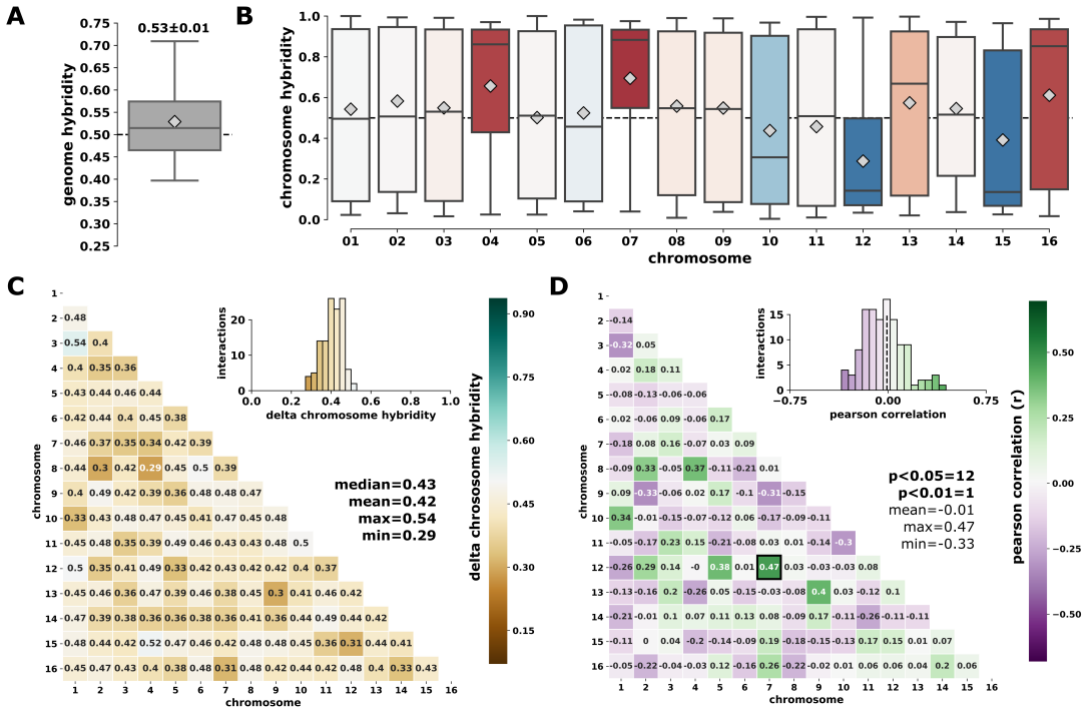

Figure S6 Analysis of hybrid genomes in environment without toxin (Kao et al. 2010) **(A)** Genome hybridity measured as the mean of the 16 chromosome hybridities within a genome. Boxplots indicate the median and interquartile range. Diamonds indicate the mean. Dashed line (0.5) indicates equal amount of the genome mapping to *S. cerevisiae*. **(B)** Chromosome hybridity for each chromosome. Chromosome hybridity is measured as the percent chromosome mapping to *S. cerevisiae*. Boxplots are colored according to where the median falls. Dashed line (0.5) indicates equal amounts of the chromosome mapping to *S. cerevisiae*. **(C)** Average change in chromosome hybridity (percent chromosome mapping to *S. cerevisiae*) between chromosomes. Inset depicts the distribution of delta chromosome hybridity for all chromosome interactions. Delta chromosome hybridity is determined for each chromosome interaction by taking the difference between the hybridity measurements for each chromosome within an F2 genome (n genomes = 36). These values are then averaged across all diploid F2 genomes, and medians and means are reported and colored accordingly. A large delta chromosome hybridity (green) suggests that the chromosomes map primarily to opposing species. A small delta chromosome hybridity (brown) suggests that the chromosomes have similar levels of hybridity. These chromosomes may come from primarily the same species or at least have similar hybridity proportions. **(D)** Pearson correlation coefficient (r) of chromosome hybridity (percent chromosome mapping to *S. cerevisiae*) between chromosomes (n genomes = 36). Inset depicts the distribution of Pearson correlation coefficients for all chromosome interactions. Positive correlations are shown in green and negative correlations are shown in purple. Statistically significant correlations ($p < 0.01$) are highlighted in black.

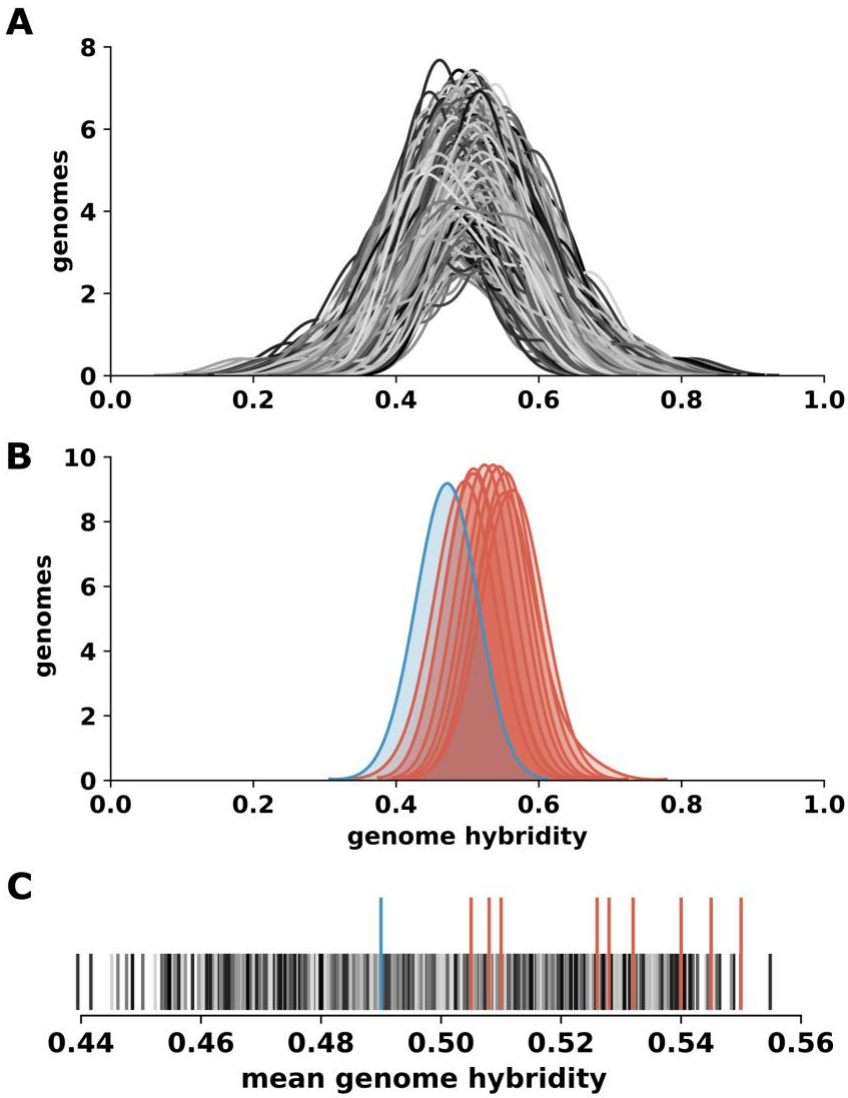

Figure S7 Genome-wide hybridity of simulated chromosomes
(A) The distribution of genome-wide hybridity in a simulated environment resulting from simulated chromosomes based on no chromosomal interactions (no epistasis). Each environment consists of 24 genomes, each consisting of 16 simulated chromosomes. Each simulated environment was repeated 10,000 times. **(B)** The distribution of genome-wide hybridity found in our experimental populations. Distributions are colored according to their overall mean species bias (red = *S. cerevisiae*, blue = *S. paradoxus*). **(C)** Mean genome-wide hybridity of our experimental populations as compared to the range of genome-wide hybridity found in simulations (grey-black lines).

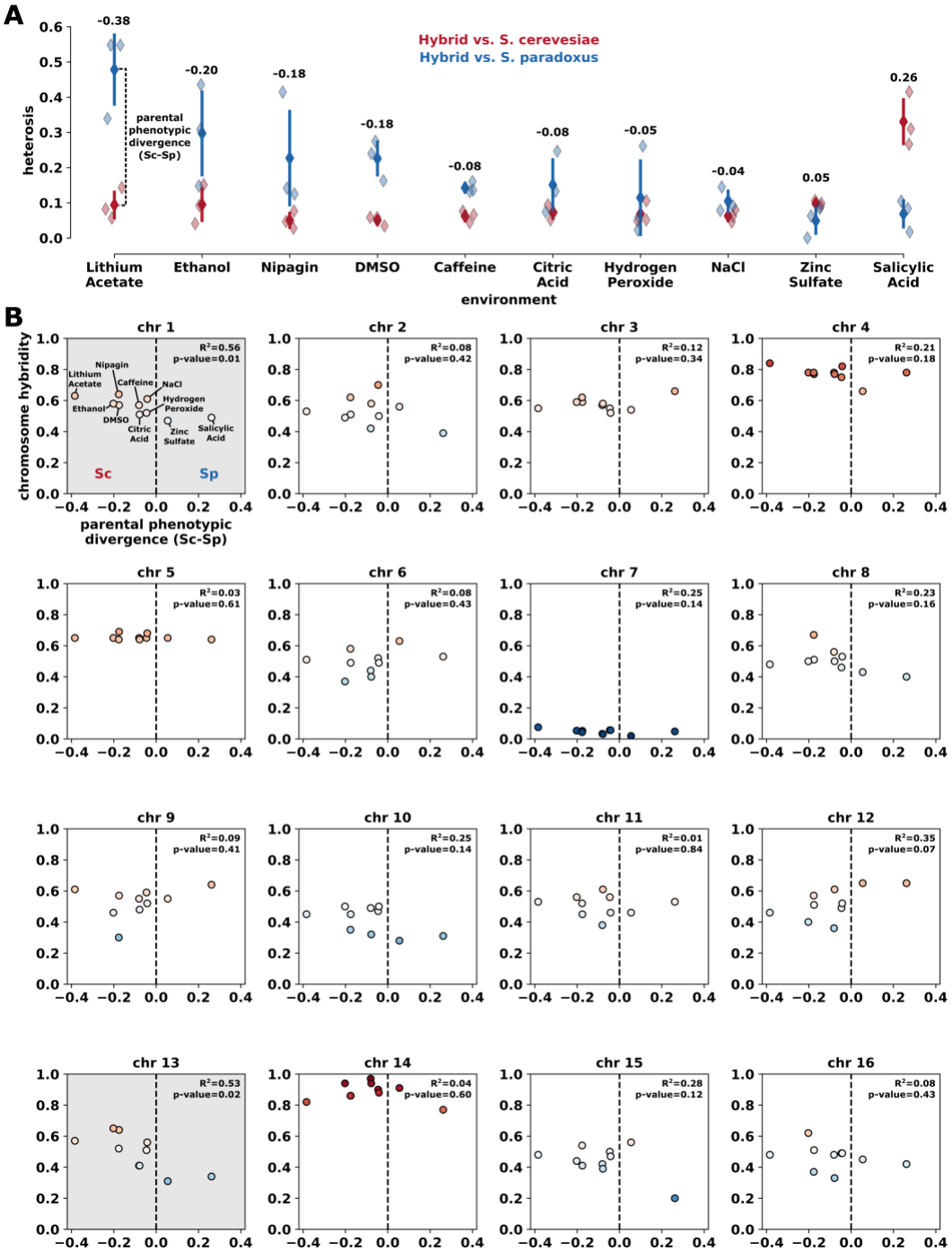

Figure S8 Analysis of parental phenotypic divergence and chromosome hybridity
(A) Heterosis for competitive growth of an interspecific hybrid in ten stressful environments (re-plotted from Bernardes et al. 2017). Solid diamonds and error bars indicate mean heterosis and standard deviation vs. *S. cerevisiae* (red) or *S. paradoxus* (blue). Parental phenotypic divergence (i.e. the difference between the competitive growth of the two parents relative to their hybrid) is shown above each environment. **(B)** Relationship between parental phenotypic divergence and chromosome hybridity. Negative parental phenotypic divergence suggests the *S. cerevisiae* parent performed better during competition with hybrid. Positive parental phenotypic divergence suggests the *S. paradoxus* parent performed better during competition with hybrid. Each environment for each chromosome is colored according to mean chromosome hybridity (Figure 2). Grey background indicates a significant correlation ($p<0.05$).

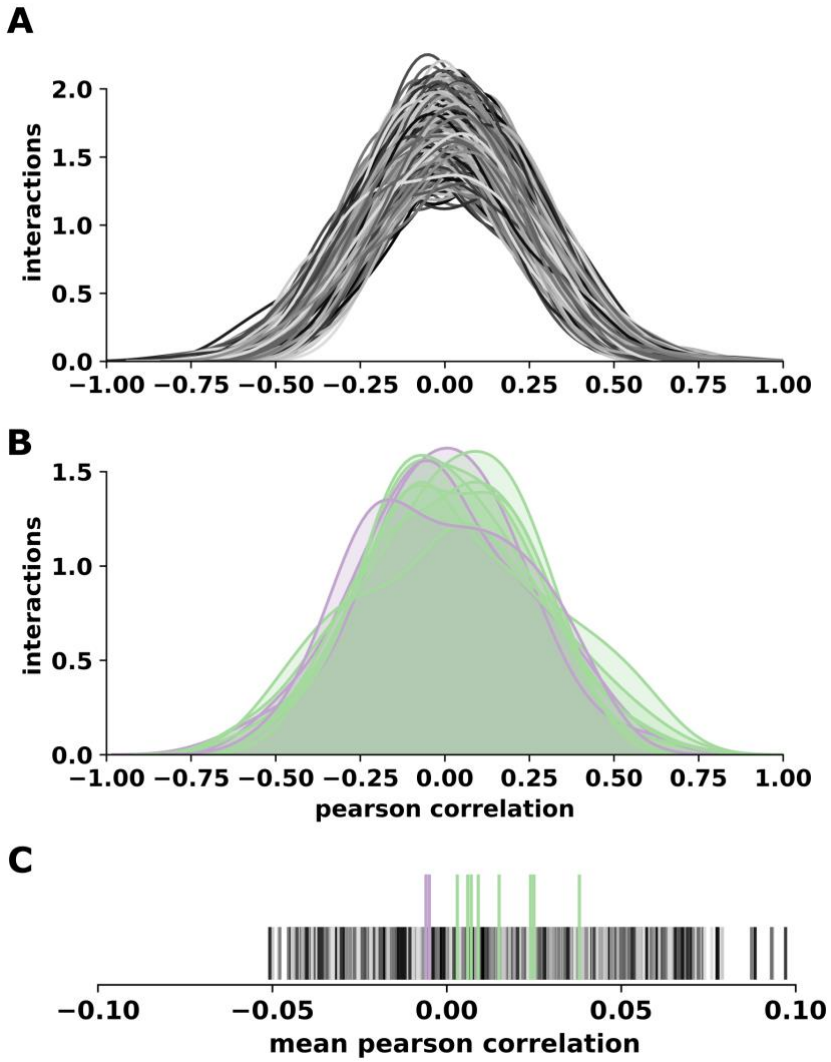

Figure S9 Pearson correlation (r) of simulated chromosomes
(A) The distribution of Pearson correlations in a simulated environment resulting from simulated chromosomes based on no chromosomal interactions (no epistasis). Each environment consists of 24 genomes, each consisting of 16 simulated chromosomes. Each simulated environment was repeated 10,000 times. **(B)** The distribution of Pearson correlations found in our experimental environments. Distributions are colored according to their mean Pearson correlation. **(C)** Mean Pearson correlation of our experimental environments as compared to the range of mean Pearson correlations found in simulations (grey-black lines).

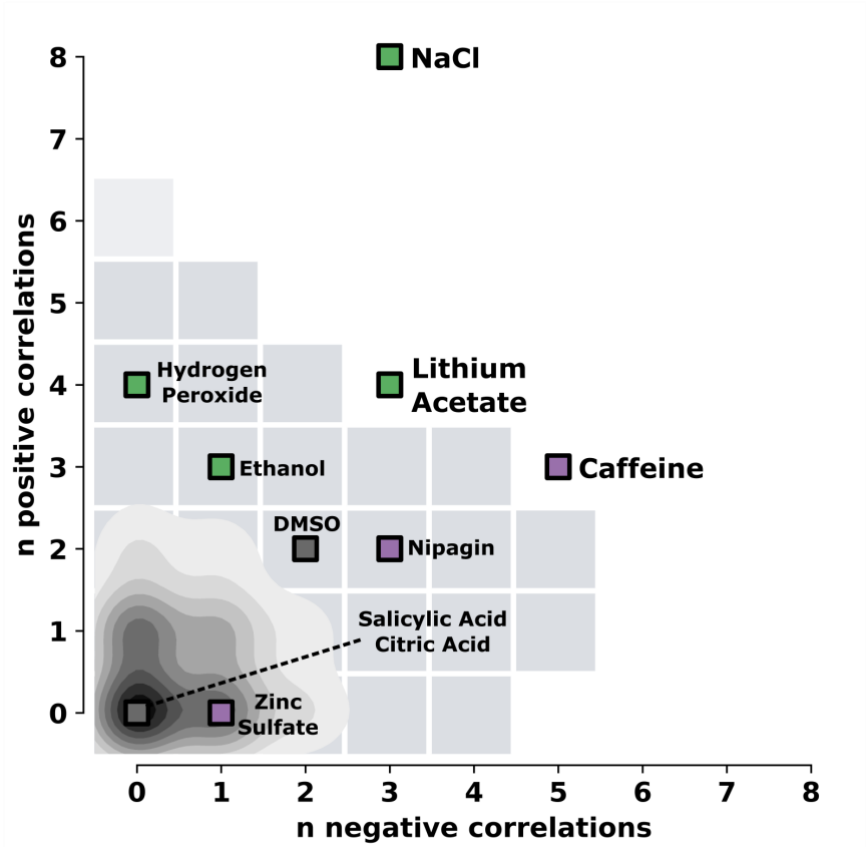

Figure S10 Distribution of significant (p < 0.01) positive and negative correlations resulting from simulated chromosomes

Each environment consists of 24 genomes, each consisting of 16 simulated chromosomes. Each simulated environment was repeated 10,000 times. A grey square indicates that at least one simulated environment resulted in that many negative and positive correlations. The density plot indicates the distribution of simulated environments. Our experimental environments are indicated with squares and labels. The color of each environment indicates the bias towards positive (green) or negative (purple) correlations.

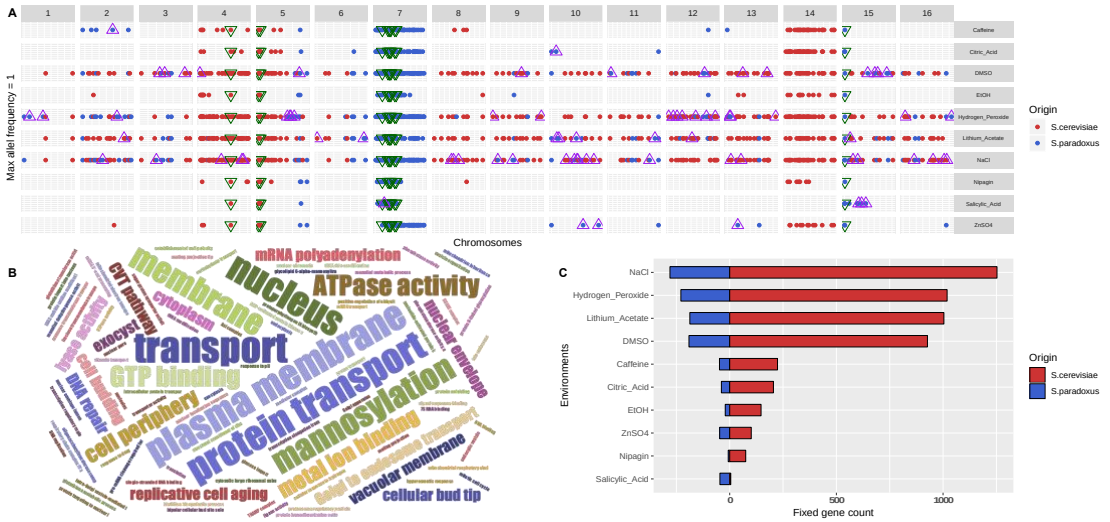

Figure S11 Genotype-by-environment interactions

(A) 10kb bins fixed for either one or the other parental allele across all genomes sampled from a given environment. All 16 chromosomes are shown. Bins fixed for the same species across all environments are marked with green triangles. Environment-specific regions fixed for the same species in all 24 genomes sampled from the same environment are marked with purple triangles. **(B)** Word cloud summarizing results of gene ontology analysis of the 1884 genes overlapping with or located within all fixed bins using DAVID (Huang, et al. 2009) **(C)** Bar chart showing the number of genes fixed for the *S. paradoxus* allele on the left and the *S. cerevisiae* allele on the right, by environment.