

Supplement for:

**Analysis of the *Saccharomyces cerevisiae* Pan-Genome Reveals a Pool of Copy
Number Variants Distributed in Diverse Yeast Strains From Differing
Industrial Environments**

Barbara Dunn¹, Chandra Richter², Daniel J. Kvitek¹, Tom Pugh² and Gavin Sherlock^{1,3}

Supplemental Figure Legends

Supplemental Figure 1. Introgression event from *S. paradoxus*. This is the same as Figure 2 in the main text, but with chromosome and contig coordinates.

Supplemental Figure 2. Corresponding *S. cerevisiae* region for strains with introgression events from *S. paradoxus* covering the *SUC2* gene.

Supplemental Figure 3. Introgression event from *S. mikatae*.

Supplemental Figure 4. aCGH results for different chromosomal regions of *S. cerevisiae* genome. A – D. Ty transposons Ty1, Ty3, Ty4, Ty5, respectively. Hierarchical clustering using the aCGH data for only the *S. cerevisiae* Ty probes was performed, with the generated dendrograms shown at top and the probe names shown along the right side (except for Ty1); note that the Ty probes were also clustered to clearly show the hybridization patterns among the strains, but the dendrograms are not shown for the probe clusters. Red and green bars indicate increased and decreased relative hybridization to the probes, respectively.

Supplemental Figure 5. aCGH data for the 2 μ plasmid.

Supplemental Figure 6. aCGH data for the YGR201C region, validated by DNA sequencing.

Supplemental Figure 7. aCGH data for the *ENA* locus, validated by DNA sequencing.

Supplemental Figure 8. aCGH data for the *HXT6/7* locus, validated by DNA sequencing.

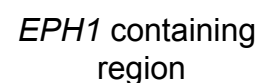
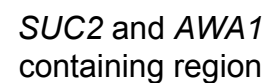
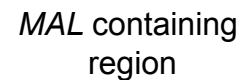
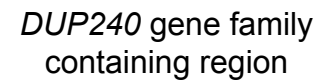
Supplemental Table Legends

Supplemental Table 1. Principal component 1 (PC1) most significant positively-correlated and most significant negatively-correlated probes, with associated genomic features.

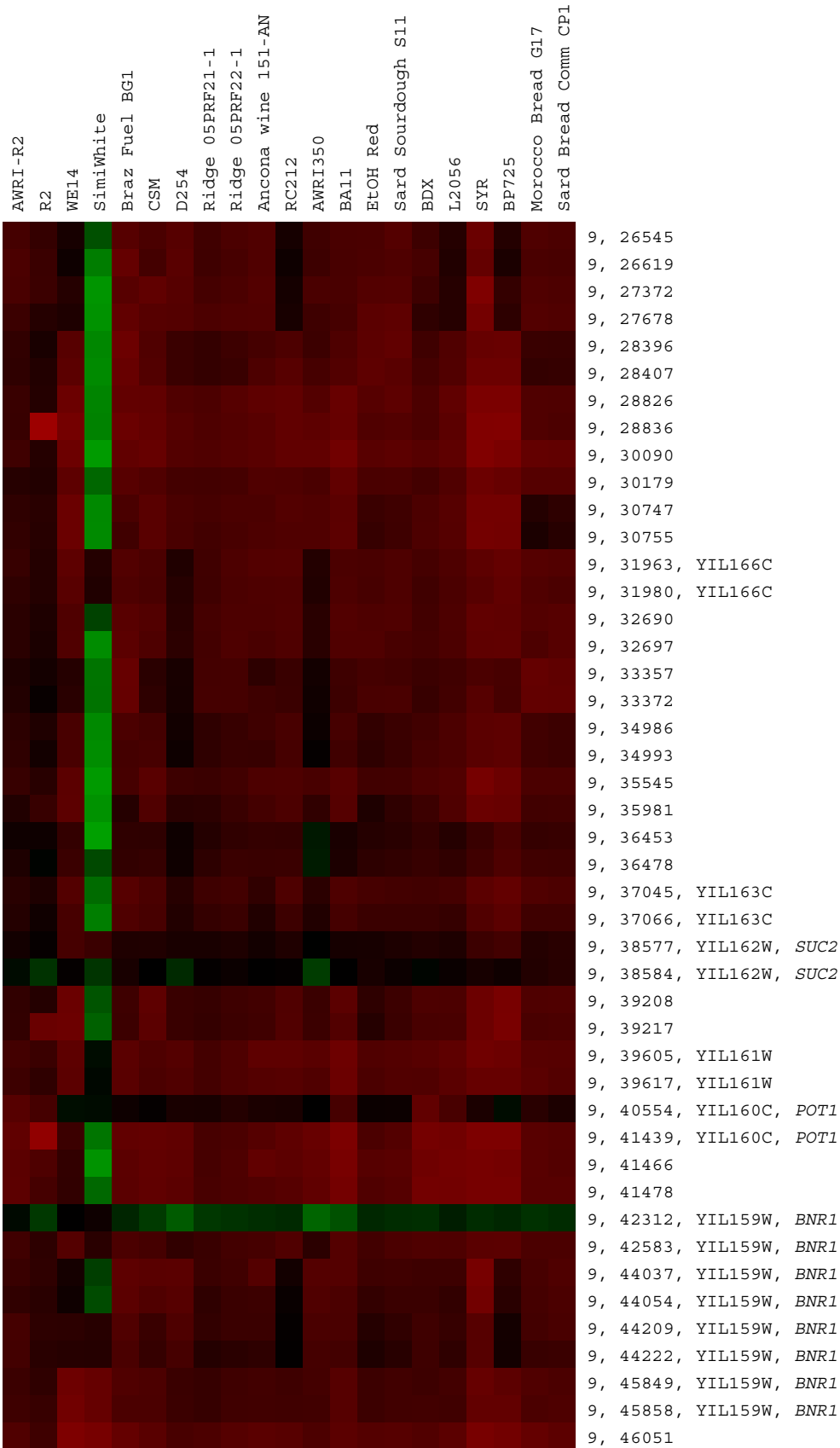
Supplemental Table 2. List of sequenced strains used for microarray design, and as common reference pool.

Supplemental Table 3. List of oligonucleotides used for validation of CNVs and introgressions.

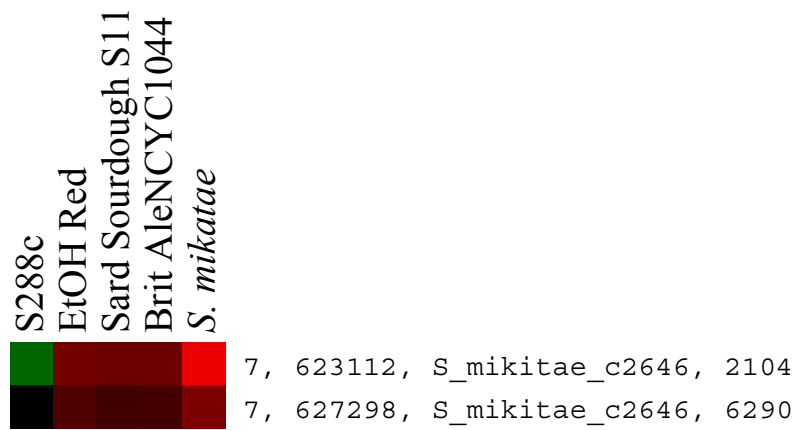
S. paradoxus
AWRI-R2
R2
WE14
Simi White
Brazilian Fuel BG1
CSM
D254
Ridge 05PRF21-1
Ridge 05PRF22-1
RC212
Ancona wine 151-AN
AWRI350
BA11
Sardinian Sourdough S11
EtOH Red



Supplemental Figure 2

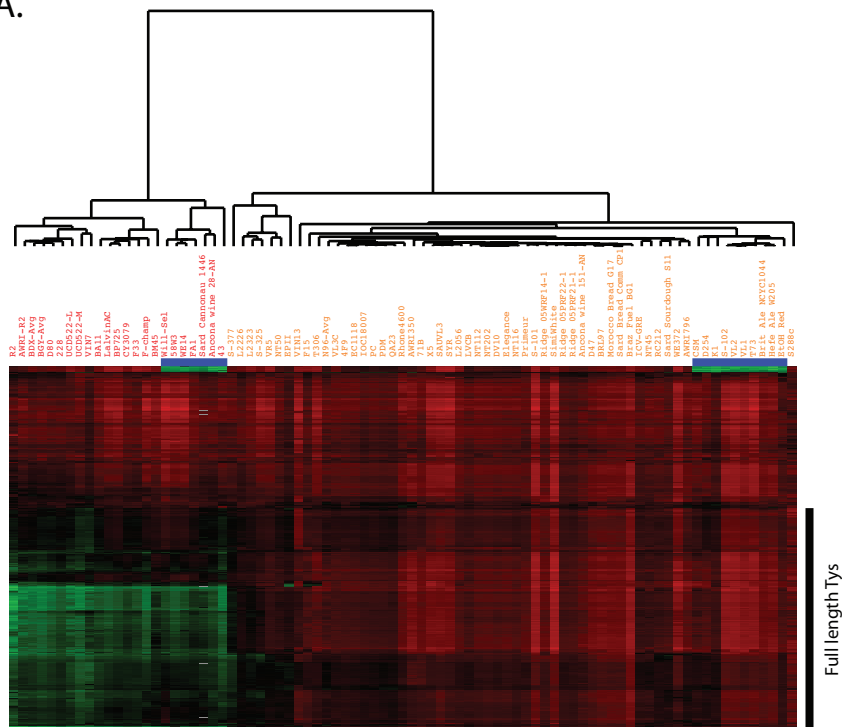


Supplemental Figure 3

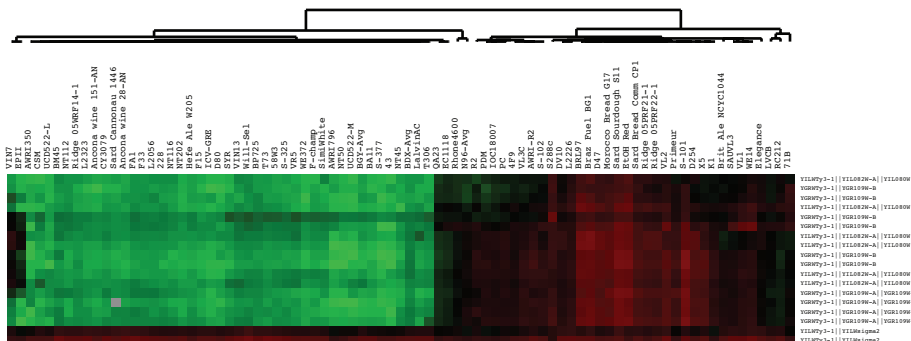


Supplemental Figure 4

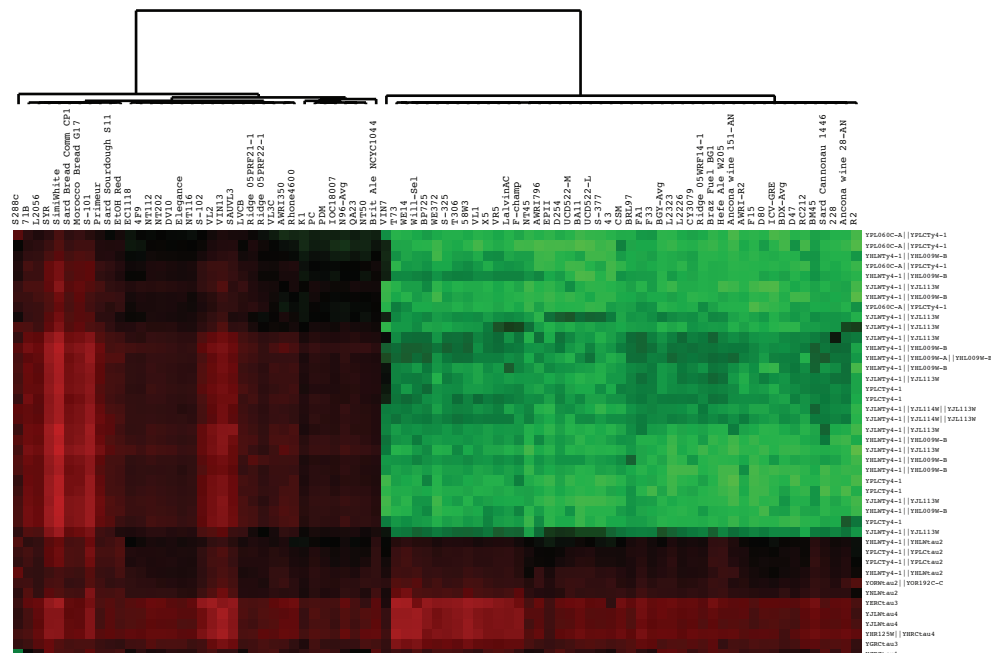
A.



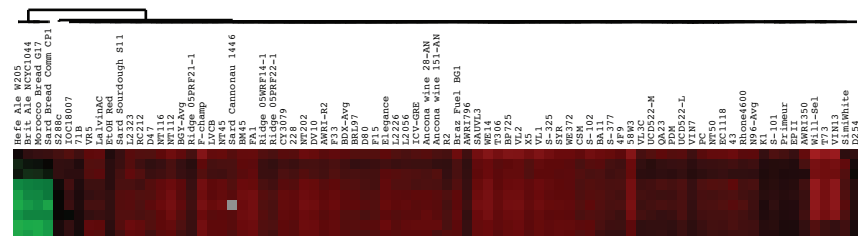
B.



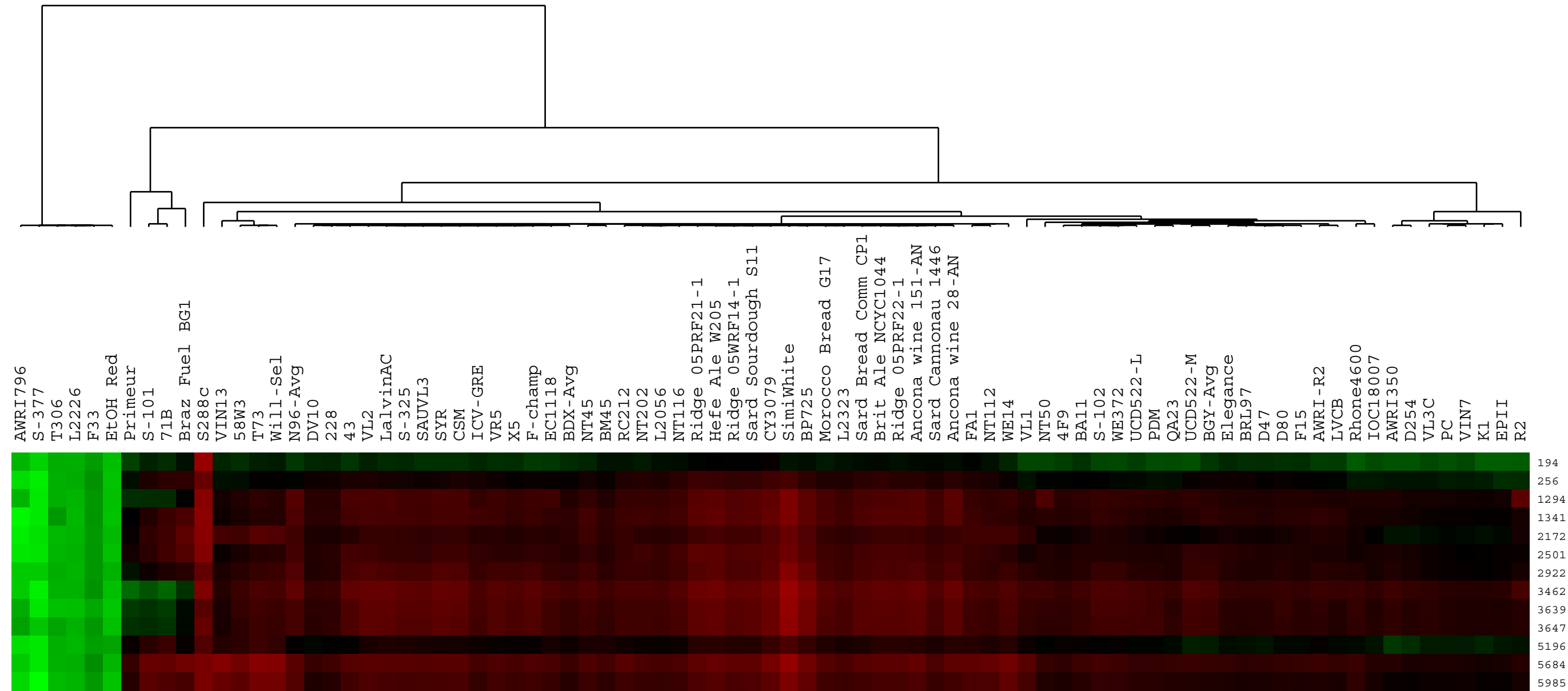
C.



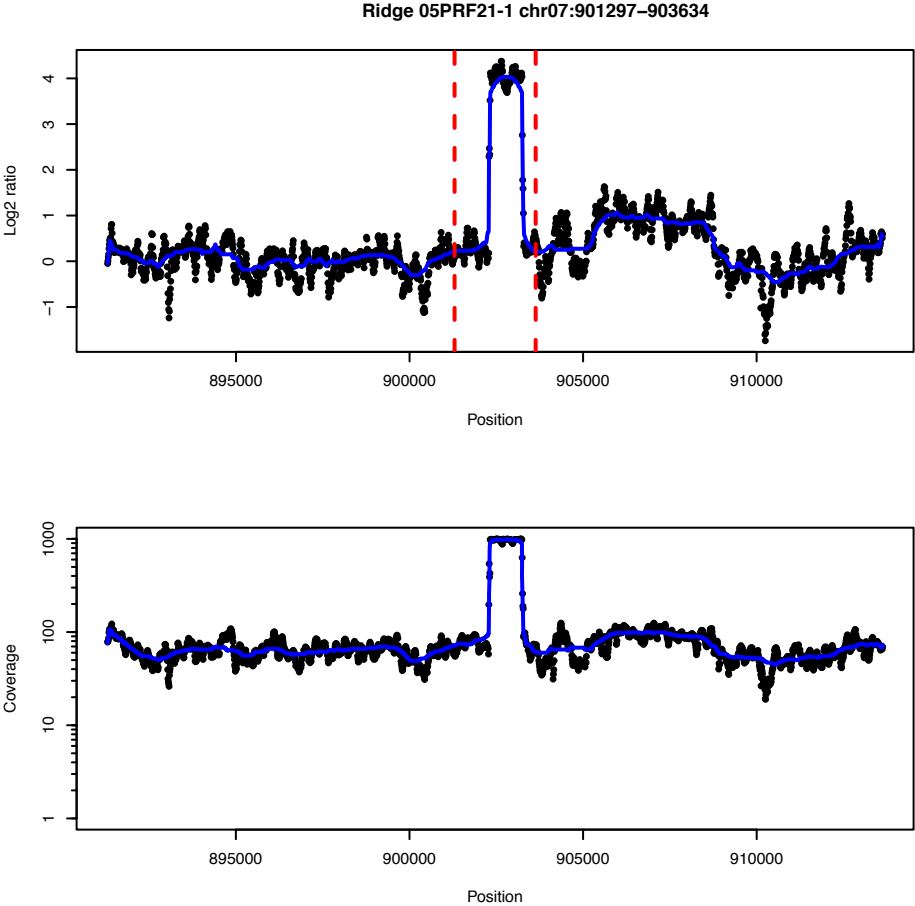
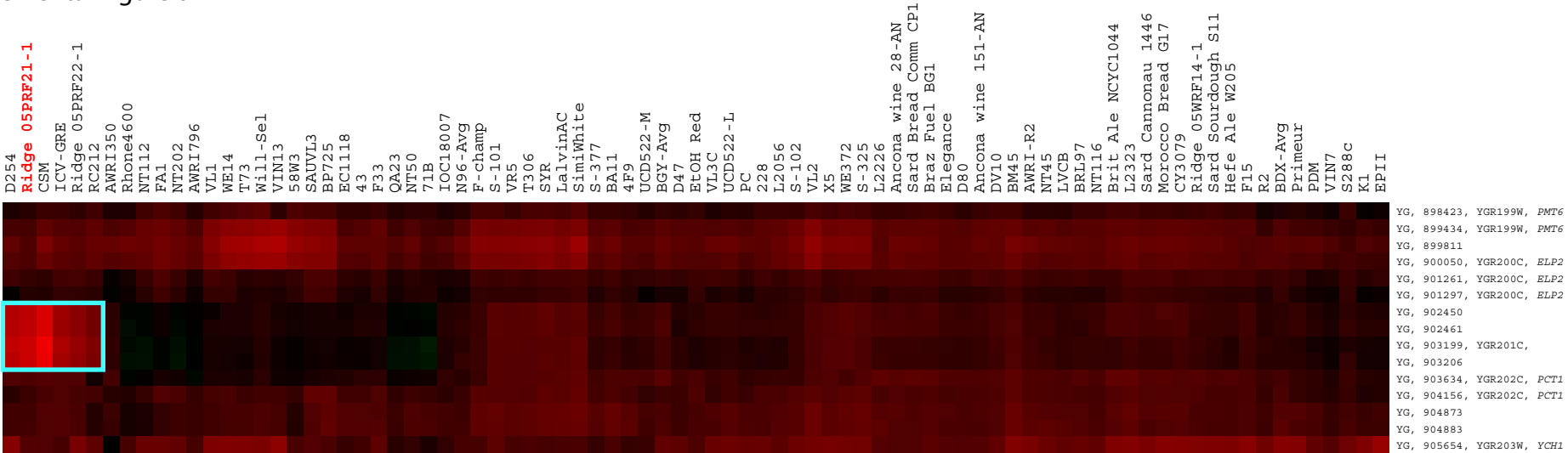
D.



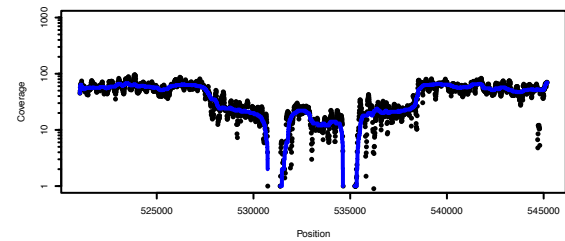
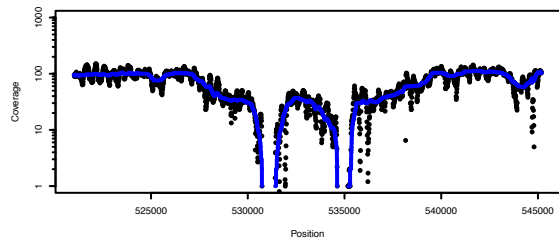
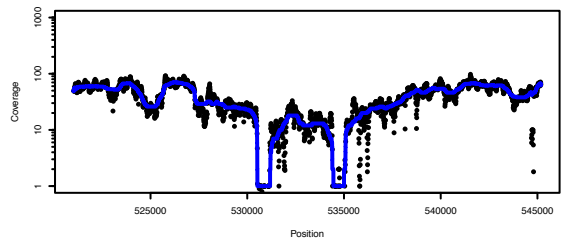
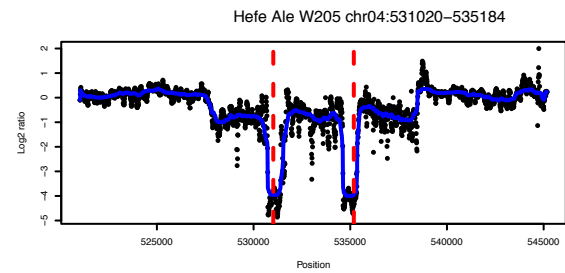
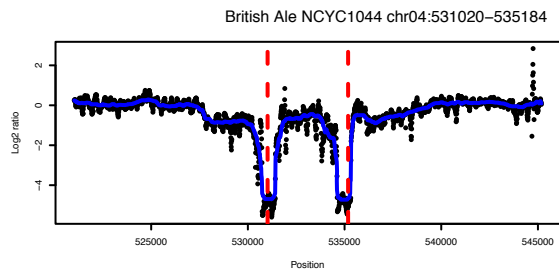
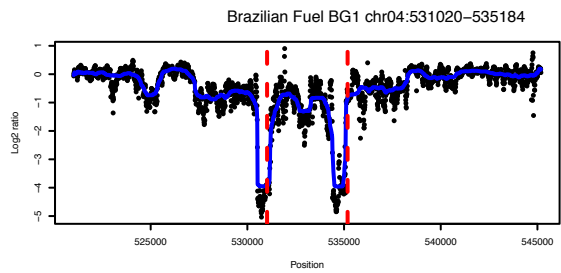
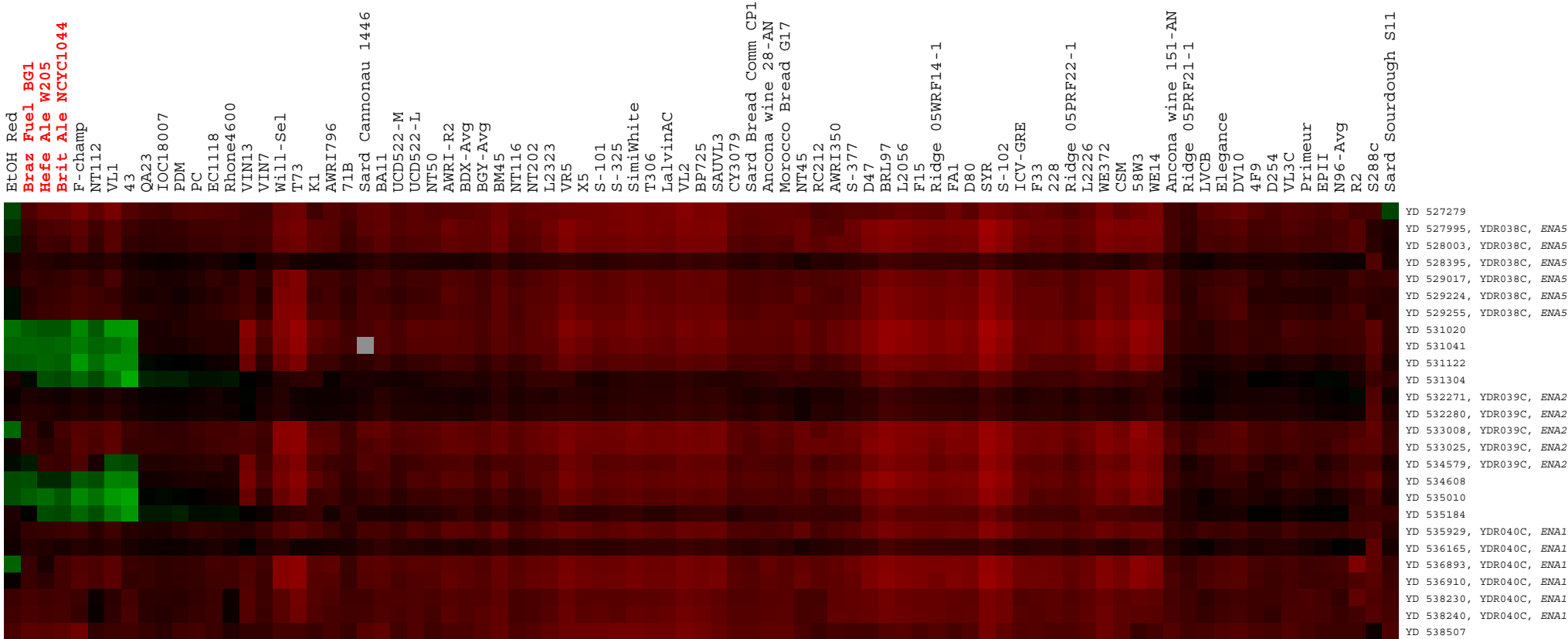
Supplemental Figure 5



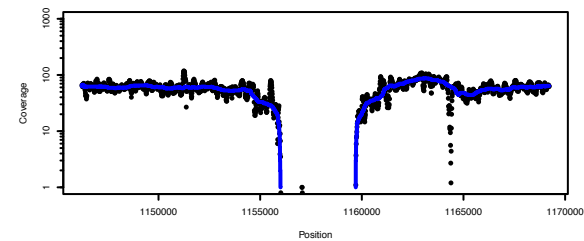
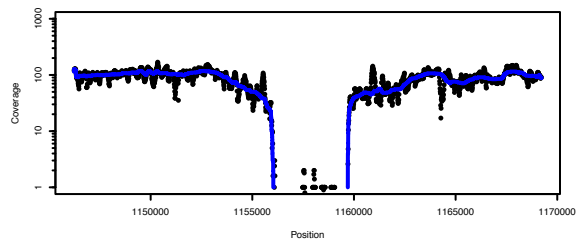
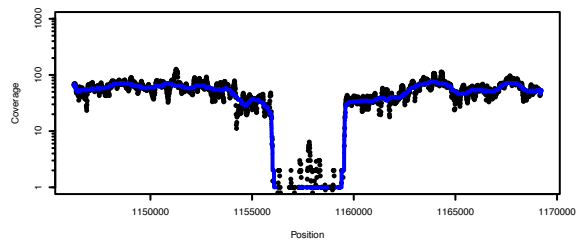
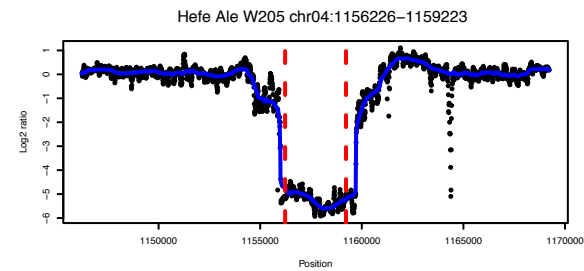
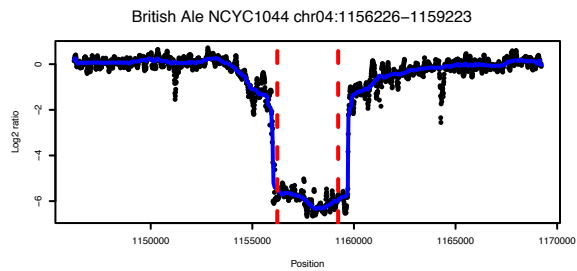
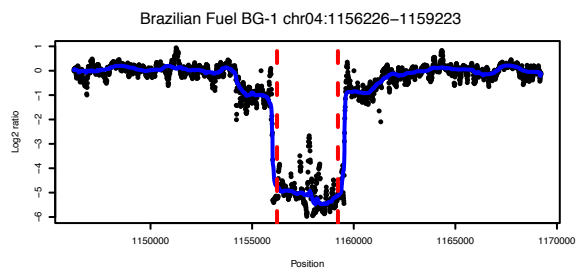
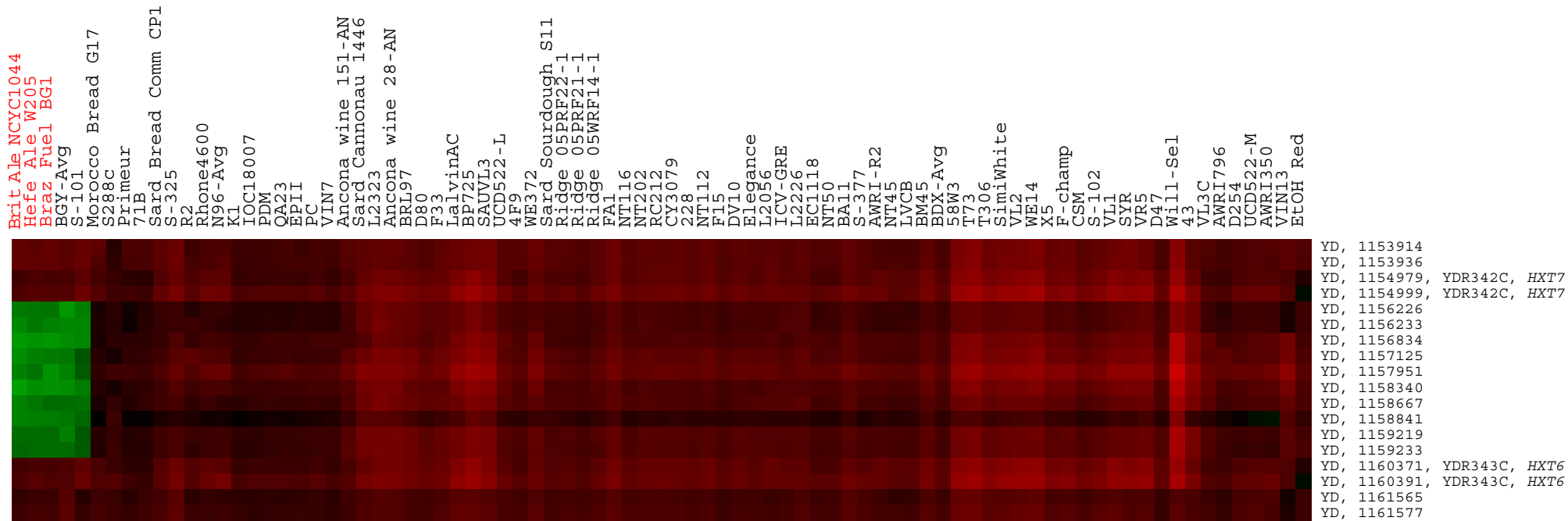
Supplemental Figure 6



Supplemental Figure 7



Supplemental Figure 8



Supplemental Table 1

Top-Most Positive Probes	PC1 loading value	Genomic features
Y55-28f16.plk bases 16 to 1023 [1-1007] 913	0.0655	Y55 upstream of <i>AGP3</i> homologue region
YF-20508	0.0638	S288c <i>AGP3/DAK2</i> region
gi 207345754 gb ABSV01000744.1 _(315..1703)_ [1-1389] 1327	0.0629	AWRI1631 upstream of <i>AGP3</i> homologue region
YF-19194	0.0615	S288c <i>AGP3/DAK2</i> region
YF-25188	0.0613	S288c <i>AGP3/DAK2</i> region
YF-16872	0.0605	S288c <i>AGP3/DAK2</i> region
YF-25506	0.0599	S288c <i>AGP3/DAK2</i> region
YF-22497	0.0586	S288c <i>AGP3/DAK2</i> region
YF-27846	0.0562	S288c <i>AGP3/DAK2</i> region
YF-28917	0.0562	S288c <i>AGP3/DAK2</i> region
YF-20754	0.0556	S288c <i>AGP3/DAK2</i> region
YF-19317	0.0554	S288c <i>AGP3/DAK2</i> region
YF-28633	0.0549	S288c <i>AGP3/DAK2</i> region
YF-20747	0.0547	S288c <i>AGP3/DAK2</i> region
YF-28908	0.0539	S288c <i>AGP3/DAK2</i> region
YF-26099	0.0533	S288c <i>AGP3/DAK2</i> region
YF-28741	0.0532	S288c <i>AGP3/DAK2</i> region
YF-22340	0.0532	S288c <i>AGP3/DAK2</i> region
YF-27855	0.0528	S288c <i>AGP3/DAK2</i> region
YF-26581	0.0510	S288c <i>AGP3/DAK2</i> region
Y55.chr10_(27105..28403)_ [1-1299] 123	0.0508	Unknown non-S288c region
Y55.chr10_(27105..28403)_ [1-1299] 131	0.0508	Unknown non-S288c region
Y55.chr10_(27105..28403)_ [1-1299] 138	0.0492	Unknown non-S288c region
gi 207345754 gb ABSV01000744.1 _(315..1703)_ [1-1389] 1186	0.0479	AWRI1631 upstream of <i>AGP3</i> homologue region
YF-26378	0.0473	S288c <i>AGP3/DAK2</i> region
gi 207345754 gb ABSV01000744.1 _(315..1703)_ [1-1389] 1175	0.0464	AWRI1631 <i>AGP3</i> homologue region
gi 207340144 gb ABSV01002472.1 _(1..3013)_ [1901-3013] 356	0.0457	AWRI1631 Unknown wine-specific protein

gi 207340144 gb ABSV01002472.1 _(1..3013)_[1901-3013]_750	0.0449	AWRI1631 Unknown wine-specific protein
gi 207340144 gb ABSV01002472.1 _(1..3013)_[1901-3013]_672	0.0433	AWRI1631 Unknown wine-specific protein
YF-20167	0.0432	S288c <i>AGP3/DAK2</i> region
YK-653061	0.0427	S288c <i>FLO10/NFT1</i> region
YJ-737134	0.0419	S288c <i>SOR1/MPH3</i> intergenic region
gi 207340128 gb ABSV01002484.1 _(1..211)_[1-211]_104	0.0416	AWRI1631 Unknown wine-specific region
YF-22846	0.0415	S288c <i>AGP3/DAK2</i> region
gi 207340128 gb ABSV01002484.1 _(1..211)_[1-211]_97	0.0407	AWRI1631 Unknown wine-specific region
YD-8470	0.0404	S288c <i>SOR2/MPH2</i> region
YJ-737926	0.0398	S288c <i>SOR1/MPH3</i> region
YD-8398	0.0397	S288c <i>SOR2/MPH2</i> region
gi 207340144 gb ABSV01002472.1 _(1..3013)_[1-2000]_1593	0.0397	AWRI1631 Unknown wine-specific protein
YD-7796	0.0396	S288c <i>SOR2/MPH2</i> region
Top-Most Negative Probes		
YJ-201002	-0.0887	Ty4
YJ-200976	-0.0874	Ty4
YP-439411	-0.0852	Ty4
YH-89178	-0.0849	Ty4
YP-439401	-0.0841	Ty4
YP-438936	-0.0827	Ty4
YH-89951	-0.0821	Ty4
YH-88401	-0.0815	Ty4
YJ-201660	-0.0808	Ty4
YH-88480	-0.0784	Ty4
YJ-199304	-0.0744	Ty4
YJ-198573	-0.0733	Ty4
YN-99121	-0.0730	Ty1
YP-437963	-0.0720	Ty4

YJ-202638	-0.0719	Ty4
YJ-198582	-0.0719	Ty4
YH-90628	-0.0718	Ty4
YH-87269	-0.0715	Ty4
YH-90665	-0.0696	Ty4
YP-438235	-0.0694	Ty4
YN-522247	-0.0693	Ty1
YP-440505	-0.0690	Ty4
YJ-10995	-0.0687	<i>VTH1</i> and <i>VTH2</i>
YI-11878	-0.0686	<i>VTH1</i> and <i>VTH2</i>
YI-11221	-0.0685	<i>VTH1</i> and <i>VTH2</i>
YN-522254	-0.0682	Ty1
YP-441317	-0.0681	Ty4
YH-86517	-0.0675	Ty4
YJ-11861	-0.0671	<i>VTH1</i> and <i>VTH2</i>
YN-99201	-0.0668	Ty1
YB-263184	-0.0661	Ty1
YJ-11205	-0.0655	<i>VTH1</i> and <i>VTH2</i>
YJ-199318	-0.0652	Ty4
YJ-202665	-0.0650	Ty4
YJ-22647	-0.0648	YJL218
YH-88085	-0.0646	Ty4
YP-440522	-0.0644	Ty4
YI-16359	-0.0642	<i>VTH1</i> and <i>VTH2</i>
YI-16309	-0.0625	<i>VTH1</i> and <i>VTH2</i>
YI-16227	-0.0621	<i>VTH1</i> and <i>VTH2</i>

Supplemental Table 2

Strain Name	Origin	Substrate	Further notes
<i>Saccharomyces cerevisiae</i> sequenced strains used for array design			
S288c	California, USA	Rotting fig (plus bakers); Lab	Haploid
RM11_1A	California, USA	Grapes	Haploid derivative of natural vineyard isolate; From L. Kruglyak
Yllc17_E5	Sauternes, France	Wine	Homozygous diploid derivative (Sanger center sequence)
BC187	Napa Valley, USA	Wine (Barrel fermentation)	spore from UCD2120; Homozygous diploid derivative (Sanger center sequence)
YPS606	Pennsylvania, USA	Oak	Homozygous diploid derivative (Sanger center sequence)
W303	Unknown	Unknown; Lab	Haploid
YPS128	Pennsylvania, USA	Oak	Homozygous diploid derivative (Sanger center sequence)
DBVPG1106	Australia	Grapes	Homozygous diploid derivative (Sanger center sequence)
DBVPG6040	Netherland	Fermenting fruit juice	ex <i>S. fructum</i> ; Homozygous diploid derivative (Sanger center sequence)
SK1	USA	Soil; Lab	Haploid
L_1374	Chile	Wine	Homozygous diploid derivative (Sanger center sequence)
L_1528	Chile	Wine	Homozygous diploid derivative (Sanger center sequence)
NCYC110	West Africa	Ginger beer from <i>Z. officinale</i>	Homozygous diploid derivative (Sanger center sequence)
NCYC361	Ireland	Beer spoilage strain from wort	Homozygous diploid derivative (Sanger center sequence)
K11	Japan	Shochu Sake strain	Awamori-1; Homozygous diploid derivative (Sanger center sequence)
Y9	Japan	Ragi (similar to sake wine)	Homozygous diploid derivative (Sanger center sequence)
Y12	Africa	Palm wine strain	Homozygous diploid derivative (Sanger center sequence)
YS2	Australia	Baker strain	Homozygous diploid derivative (Sanger center sequence)
YS4	Netherland	Baker strain	Homozygous diploid derivative (Sanger center sequence)
YS9	Singapore	Baker strain	Homozygous diploid derivative (Sanger center sequence)
UWOPS83_787_3	Bahamas	Fruit, <i>Opuntia stricta</i>	Homozygous diploid derivative (Sanger center sequence)
UWOPS87_2421	Hawaii	Cladode, <i>Opuntia megacantha</i>	Homozygous diploid derivative (Sanger center sequence)
UWOPS03_461_4	Malaysia	Nectar, Bertam palm	Homozygous diploid derivative (Sanger center sequence)
UWOPS05_217_3	Malaysia	Nectar, Bertam palm	Homozygous diploid derivative (Sanger center sequence)
UWOPS05_227_2	Malaysia	Trigona, Bertam palm	Homozygous diploid derivative (Sanger center sequence)
322134S	RVI, Newcastle UK	Clinical isolate (Throat-sputum)	Homozygous diploid derivative (Sanger center sequence)
378604X	RVI, Newcastle UK	Clinical isolate (Sputum)	Homozygous diploid derivative (Sanger center sequence)
273614N	RVI, Newcastle UK	Clinical isolate (Fecal)	Homozygous diploid derivative (Sanger center sequence)
DBVPG6765	Unknown	Unknown	ex <i>S. boulardii</i> ; Homozygous diploid derivative (Sanger center sequence)
YJM978	Bergamo, Italy	Clinical isolate (Vaginal)	Homozygous diploid derivative (Sanger center sequence)
YJM981	Bergamo, Italy	Clinical isolate (Vaginal)	Homozygous diploid derivative (Sanger center sequence)
YJM975	Bergamo, Italy	Clinical isolate (Vaginal)	Homozygous diploid derivative (Sanger center sequence)
DBVPG6044	West Africa	Bili wine	ex <i>S. manginii</i> ; Homozygous diploid derivative (Sanger center sequence)
DBVPG1788	Finland	Soil	Homozygous diploid derivative (Sanger center sequence)
DBVPG1373	Netherland	soil	Homozygous diploid derivative (Sanger center sequence)
DBVPG1853	Ethiopia	White Tecc	Homozygous diploid derivative (Sanger center sequence)
Y55	France	Wine; Lab	Haploid
YJM789		Lung of AIDS pneumonia patient	From Ron Davis lab; Haploid derivative of clinical isolate
M22	Italy	Wine	From Justin Fay; probable heterozygous diploid
YPS163	Pennsylvania, USA	Oak	From Justin Fay; probable heterozygous diploid
AWRI1631	Australia	Commercial wine strain	Haploid deriv of N96 (related to EC1118) commercial wine strain
Other <i>Saccharomyces</i> species			
<i>S. kudriavzevii</i>	IFO1802		From Gianni Liti and Ed Louis
<i>S. mikatae</i>	IFO1815		From Gianni Liti and Ed Louis
<i>S. paradoxus</i>	CBS432	Russia, Oak	
<i>S. uvarum</i>	CBS7001 (DBVPG6299)	<i>S. uvarum</i>	
<i>S. castellii</i>	CBS4309 (DBVPG6298)	<i>S. castellii</i> , FM476	From Mark Johnston
<i>S. kluyveri</i>		<i>S. kluyveri</i> , FM479	From Mark Johnston

Supplemental Table 3

Rearrangement name/type	Primer 1 name (species, chromosome, coordinate)	Primer 1 sequence	Primer 2 name (species, chromosome ,coordinates)	Primer 2 sequence	Strain(s) that exhibit rearrangement by PCR test
YGR201C amplification	ScerChr7-903948	CTTGGTTGAAGAAAACGAACTG	ScerChr7-901570	CTCCACAATCTAATATAACGGTCCT	Ridge05PRF21-1
<i>HXT6/7</i> deletion	ScerChr4-1153512	AATGTAATTTTCATCAAATCCACG	ScerChr4-1161997	AGTGACAACGAGCACGGGT	MoroccoBreadG17,BrazFuelBG1,BritAleNCYC1044,HefeAleW205,BGY,S101
Introgression of <i>S. paradoxus</i> MAL region	ScerChr7-1061853	CACACGAAGTAGGACAACAAGCTG	SparCh7-1009470	GGATTGCAACCTATGGAAGCGTTCAA	BrazFuelBG1
Introgression of <i>S. paradoxus</i> STL1-PAD1-FDC1 region	ScerChr4-1507860	CAGAATTGCAGTTCAGGAGTAGTCACA	SparChr4-1486362	CTGTAGATTTTGAGATTACCTCTGGA	BrazFuelBG1