Supplements

1. **Meersman E, Steensels J, Mathawan M, Wittocx PJ, Saels V, Struyf N, Bernaert H, Vrancken G, Verstrepen KJ.** 2013. Detailed analysis of the microbial population in Malaysian spontaneous cocoa pulp fermentations reveals a core and variable microbiota. PLOS ONE **8**:e81559. doi:10.1371/journal.pone.0081559.
2. **Young TW, Yagiu M.** 1978. A comparison of the killer character in different yeasts and its classification. Antonie Van Leeuwenhoek **44**:59-77.
3. **van Wyk H, Divol B.** 2010. Recovery of endo-polygalacturonase activity in wine yeast and its effect on wine aroma. FEMS Yeast Res. **10**:58-71.
4. **Warringer J, Zorgo E, Cubillos FA, Zia A, Gjuvsland A, Simpson JT, Forsmark A, Durbin R, Omholt SW, Louis EJ, Liti G, Moses A, Blomberg A.** 2011. Trait variation in yeast is defined by population history. PLoS Genet. **7**:e1002111. doi:10.1371/journal.pgen.1002111.

Fig. S1. Intraspecific diversity of the *Saccharomyces cerevisiae* population in two spontaneous Malaysian cocoa pulp fermentations. Box fermentation 1 (B1, October 2011), heap fermentation 1 (H1, October 2011), as described in (1). Clustering of the strains based on the interdelta fingerprints of the different *S. cerevisiae* isolates (B1: 72 isolates, H1: 108 isolates). For similarities lower than 85%, a high-resolution fingerprint was used for confirmation (see Materials and Methods for details). Clusters consisting of the same (or very closely related) strain(s) are represented using a color code. Numbers next to branches indicate bootstrap values (%), based on 500 bootstrap simulations.

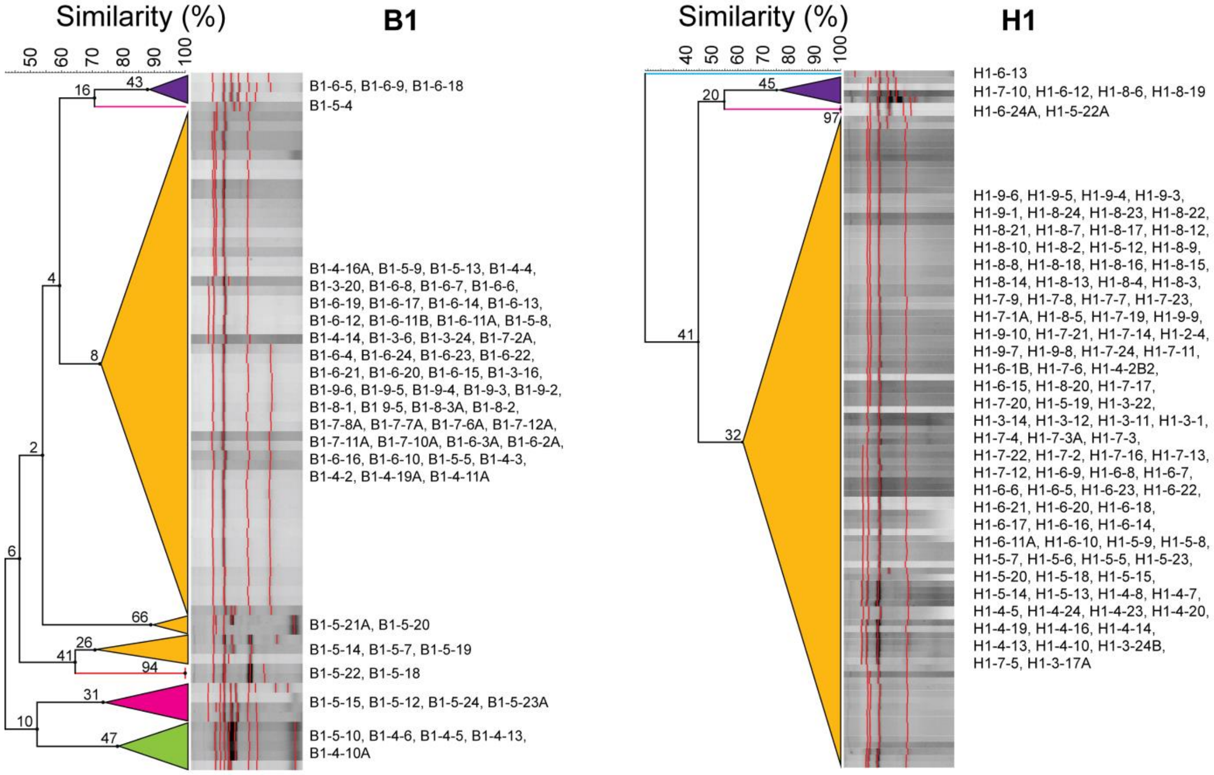

Fig. S2. Interdelta profiles of first-generation hybrids generated using mass mating. The thirty-two strains selected after mass mating (H1-H32) were identified as 13 outbred hybrids, 3 Y115 inbred strains, 15 Y927 inbreds and 1 Y927 haploid using mating type PCR and interdelta analysis. A representative selection of these different groups is shown as an illustration. Strains selected for lab- and pilot-scale fermentations are indicated in orange.

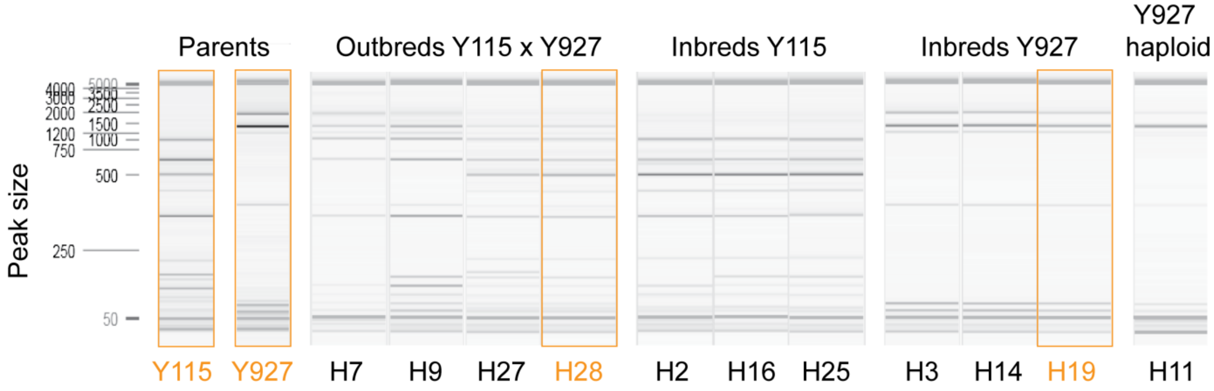

Fig. S3. Bioscreen experiments to check influence of Hygromycin B-resistance marker on growth rate. Six biological replicates were tested per strain (see Materials and Methods). A: 37°C (left), B: 41°C (right). Error bars indicate 95% confidence intervals.

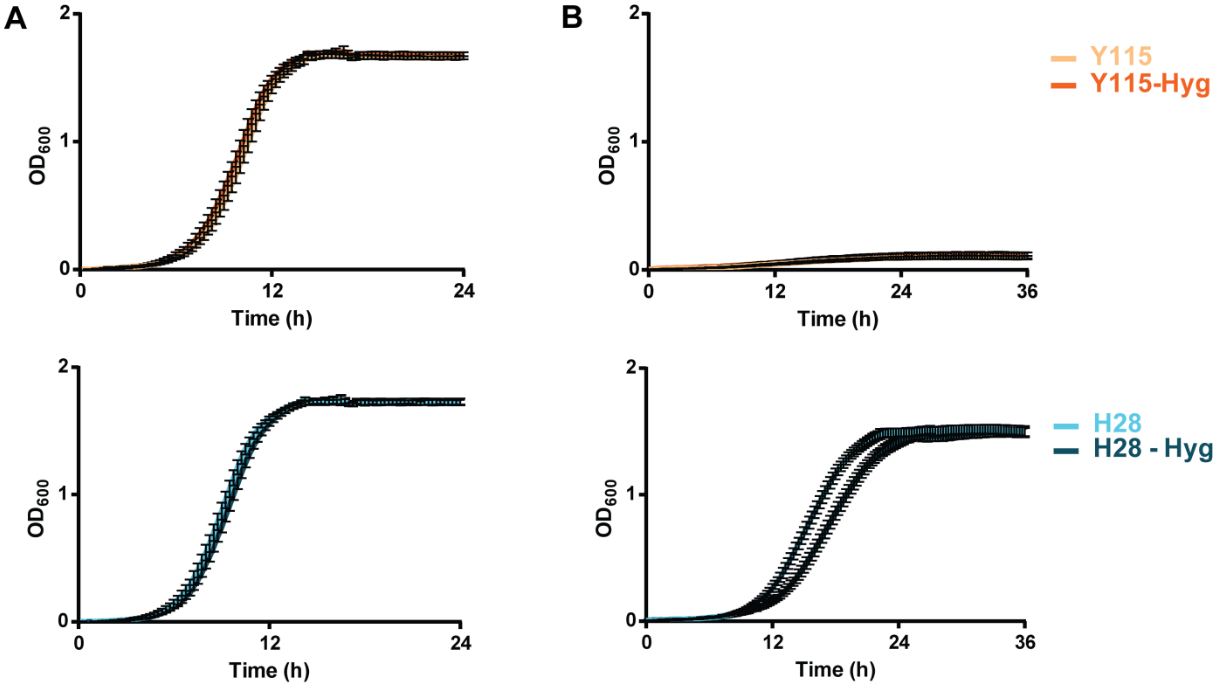

Fig. S4. Competition experiments in cocoa pulp with a Hygromycin (Hyg)-tagged transformant of the hybrid H28 and H37. Error bars represent the 95% confidence intervals, calculated based on three biological replicates. See Materials and Methods for details.

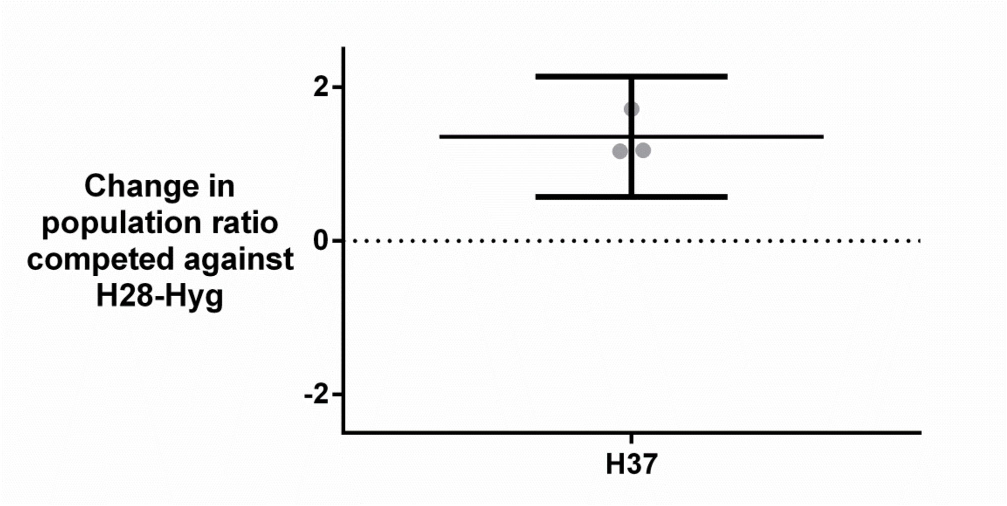

Fig. S5. Yeast population dynamics of inoculated fermentations in Malaysia (2013). All yeasts were tested in two independent fermentation assays, graphs displayed here are the "B"-duplicates from the inoculated fermentations shown in Fig. 4. Temperature (right axis) is indicated as a dashed black line. The dashed arrows indicate when the yeast cell count rises to or drops below the detection limit. Note that the scale for the total yeast count (left Y-axis) is logarithmic, whereas individual yeast species in the population at a given time point (indicated by the colors) are presented as the relative fraction of the total population on a linear scale. Absolute concentrations for individual species are given in Table S7. Percentages of identity are reported when <100% identity with type strain sequences (see Materials and Methods for details). Turning of the fermenting cocoa beans was carried out after 48 h of fermentation (solid vertical arrows). *C.: Candida, H.: Hanseniaspora, P.: Pichia, S.: Saccharomyces, T.: Torulaspora.*

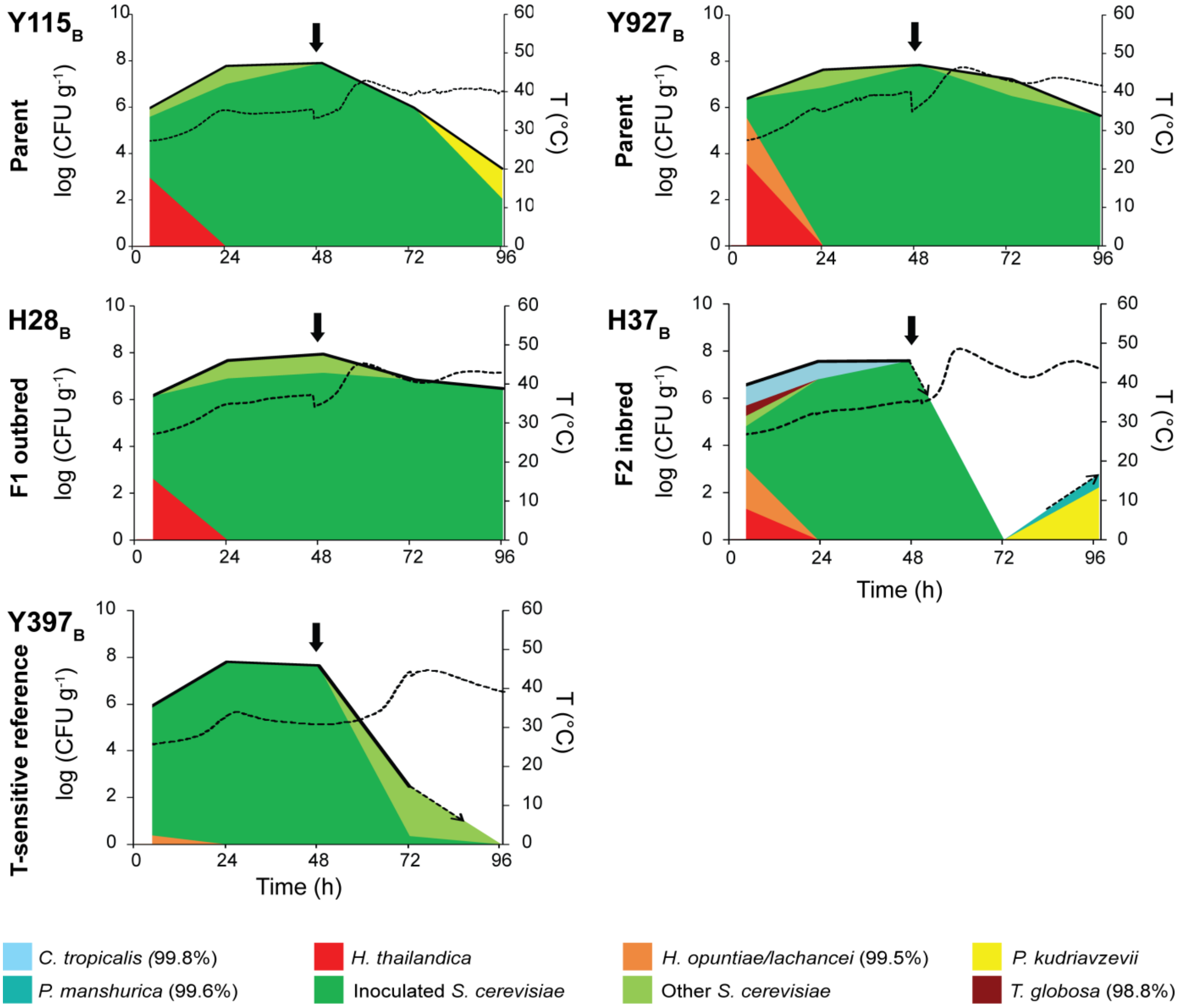

Table S1. Overview of 318 different industrial strains used in this study.

Industry	Number of strains
Beer (ale)	121
Wine	71
Feral (wild) isolates	53
Beer (lager)	31
Sake	14
Spirits	11
Bakery	10
Bio-ethanol	7

Table S2. Killer activity of first-generation hybrids resulting from mass mating Y115 and Y927.

Strain code	Strain genotype	Killer activity
Y115	Wild-type	-
Y927	Wild-type	++
H1	Hybrid (Inbred Y927)	++
H2	Hybrid (Inbred Y115)	-
H3	Hybrid (Inbred Y927)	++
H4	Hybrid (Inbred Y927)	++
H5	Hybrid (Outbred Y115 x Y927)	-
H6	Hybrid (Inbred Y927)	++
H7	Hybrid (Outbred Y115 x Y927)	-
H8	Hybrid (Inbred Y927)	++
H9	Hybrid (Outbred Y115 x Y927)	-
H10	Hybrid (Outbred Y115 x Y927)	-
H11	Haploid Y927	++
H12	Hybrid (Outbred Y115 x Y927)	-
H13	Hybrid (Inbred Y927)	++
H14	Hybrid (Inbred Y927)	++
H15	Hybrid (Outbred Y115 x Y927)	-
H16	Hybrid (Inbred Y115)	-
H17	Hybrid (Inbred Y927)	++
H18	Hybrid (Inbred Y927)	++
H19	Hybrid (Inbred Y927)	++
H20	Hybrid (Outbred Y115 x Y927)	-
H21	Hybrid (Inbred Y927)	++
H22	Hybrid (Outbred Y115 x Y927)	-
H23	Hybrid (Outbred Y115 x Y927)	-
H24	Hybrid (Outbred Y115 x Y927)	-
H25	Hybrid (Inbred Y115)	-
H26	Hybrid (Outbred Y115 x Y927)	-
H27	Hybrid (Outbred Y115 x Y927)	-
H28	Hybrid (Outbred Y115 x Y927)	-
H29	Hybrid (Inbred Y927)	++
H30	Hybrid (Inbred Y927)	++
H31	Hybrid (Inbred Y927)	++
H32	Hybrid (Inbred Y927)	++

-: no activity, +: less than reference, ++: equal to reference, +++: more than reference NCYC1001 (killer type K2, (2)). Diploids resulting from mating two cells of opposite mating types of the same parent, either germinating spores or haploid segregants, were classified as inbreds.

Table S3. Statistical differences between parental strains and hybrids during competition experiments in cocoa pulp.

	Y115-Hyg	Y927
Y115-Hyg	/	0.0261 (*)
Y927	0.0261 (*)	/
H19	0.7431 (ns)	0.0057 (**)
H28	< 0.0001 (****)	< 0.0001 (****)
H37	< 0.0001 (****)	< 0.0001 (****)

P-values were calculated using Dunnett's multiple comparisons test. *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, ****: $p \leq 0.0001$, ns: not significant. The comparison between Y115-Hyg and H19 is biologically not relevant, as the inbred H19 is only related to Y927. This value is therefore only given as an illustration.

Table S4. Overview of selected yeast strains for inoculated cocoa pulp fermentations in Malaysia.

	Killer activity	Polygalacturonase activity	Temperature tolerance
H19	++	+	+++
H28	-	-	+++
H37	-	-	+++
Y115	-	+	+++
Y397	-	-	+
Y927	++	+	+++

-: no activity, +: less than reference, ++: equal to reference, +++: more than reference. Reference for killer activity: NCYC1001 (killer type K2) (2), polygalacturonase activity: L2323 (3) and temperature tolerance: W303 (4).

Table S5. Temperature and pH of cocoa pulp in spontaneous and inoculated cocoa pulp fermentations in Malaysia (2013).

	T (°C)				pH			
	T_i	T_f	T_{min}	T_{max}	pH_i	pH_f	pH_{min}	pH_{max}
S_A	25.9	43.9	25.7	49.3	3.8	4.1	3.6	4.1
S_B	25.6	42.3	25.5	45.8	3.8	4.3	3.7	4.3
$H19_A$	26.1	42.8	26.0	48.1	3.5	4.1	3.5	4.5
$H28_A$	27.3	41.4	27.2	46.1	3.8	5.4	3.8	5.5
$H28_B$	27.2	42.9	27.1	45.3	3.9	4.0	3.9	4.8
$H37_A$	26.6	43.2	26.6	45.8	3.6	3.9	3.6	4.3
$H37_B$	26.8	44.0	26.8	48.6	3.7	3.7	3.7	4.3
$Y115_A$	27.6	41.7	27.5	45.9	3.8	4.3	3.7	4.7
$Y115_B$	27.3	40.1	27.3	43.0	3.7	6.0	3.7	6.1
$Y397_A$	26.5	43.1	26.4	49.1	3.7	4.2	3.6	4.2
$Y397_B$	25.8	39.1	25.7	44.7	3.5	4.2	3.5	4.5
$Y927_A$	30.0	41.2	27.3	45.6	3.7	4.7	3.6	4.7
$Y927_B$	27.5	41.7	27.5	46.4	3.6	4.7	3.6	4.7

Different parameters are listed for each fermentation: pH_i = initial pH, pH_f = final pH, pH_{min} = minimal pH, pH_{max} = maximal pH, T_i = initial temperature, T_f = final temperature, T_{min} = minimal temperature, T_{max} = maximal temperature. Spontaneous fermentations (S_A and S_B) and inoculated fermentations (six yeast strains in duplicate). For H19, however, only one replicate was available.

Table S6. Overview of all detected lactic and acetic acid bacterial species in the different cocoa pulp fermentations in Malaysia.

		Time point (h)				
	N° isolates	4	24	48	72	96
Lactic acid bacteria						
Inoculated						
Lactobacillus fermentum (99.6%)	22	x	x	x	x	x
Leuconostoc (pseudo)mesenteroides (98.2%)	19	x	x		x	x
Lactobacillus plantarum/paraplantarum/pentosus/fabifermentans (99.6%)	17		x	x	x	
Fructobacillus pseudoficulneus	11	x	x			
Lactobacillus amylovorus/ultunensis/gallinarum (99.6%)	9	x	x			x
Lactobacillus brevis (99.5%)	5			x	x	
Fructobacillus tropaeoli (99.6%)	4		x	x	x	
Lactobacillus mali/aquaticus/hordei (99.3%)	2			x	x	
Lactobacillus rhamnosus (99.1%)	1				x	
Spontaneous						
Lactobacillus fermentum (99.6%)	17	x	x	x	x	x
Lactobacillus plantarum/paraplantarum/pentosus/fabifermentans (99.6%)	4	x	x	x		
Lactobacillus brevis (99.5%)	3		x	x		
Leuconostoc (pseudo)mesenteroides (98.2%)	2		x			
Fructobacillus tropaeoli (99.6%)	1			x		
Acetic acid bacteria						
Inoculated						
Acetobacter pasteurianus (99.9%)	82	x	x	x	x	x
Acetobacter tropicalis/senegalensis (99.1%)	15	x	x	x		x
Acetobacter ghanensis	1			x		
Spontaneous						
Acetobacter pasteurianus (99.9%)	22	x	x	x	x	x
Acetobacter tropicalis/senegalensis (99.1%)	5	x	x	x		

All fermentations were performed in October-November 2013 in Malaysia. Data is collected from 11 inoculated and 2 spontaneous fermentations. Bacterial isolates were identified by $(GTG)_5$ rep-PCR fingerprinting, clustering and 16S rDNA sequencing of several representatives of the different clusters. Obtained sequences were compared to type strain sequences and species were identified based on the highest identity (%) of the BLAST output. Percentages of identity are reported when <100% identity. Multiple identities are given in the case of ambiguous identification.

Table S7. Population sizes of four most prevalent yeast species during spontaneous and inoculated cocoa pulp fermentations.

		log (CFU g^{-1})												
	Time (h)	H19$_A$	H28$_A$	H28$_B$	H37$_A$	H37$_B$	S$_A$	S$_B$	Y115$_A$	Y115$_B$	Y927$_A$	Y927$_B$	Y397$_A$	Y397$_B$
Hanseniaspora thailandica	4	4.43	1.97	2.68	2.59	1.31	2.42	5.50	1.14	2.98	2.71	3.58	2.65	-
	24	-	-	-	-	-	3.21	1.24	-	-	2.32	-	1.47	-
	48	-	-	-	-	-	-	-	-	-	-	-	-	-
	72	-	-	-	-	-	-	-	-	-	-	-	-	-
	96	-	-	-	-	-	-	-	-	-	-	-	-	-
Hanseniaspora opuntiae/lachancei	4	-	1.97	-	1.03	1.75	0.35	-	0.76	-	0.78	1.99	-	0.39
	24	-	-	-	-	-	-	1.24	-	-	-	-	-	-
	48	-	-	-	-	-	-	-	-	-	-	-	-	-
	72	-	-	-	-	-	-	-	-	-	-	-	-	-
	96	-	-	-	-	-	-	-	-	-	-	-	-	-
Saccharomyces cerevisiae (inoculated)	4	2.01	1.97	3.45	2.07	1.75	-	-	3.42	2.60	2.71	0.80	3.03	5.50
	24	7.54	7.77	6.90	6.83	6.81	-	-	6.87	7.00	4.73	6.87	5.14	7.81
	48	8.12	7.48	7.15	7.76	7.59	-	-	5.65	7.90	7.71	7.83	6.51	7.66
	72	4.98	6.41	6.85	5.18	-	-	-	4.11	5.98	5.97	6.49	-	0.35
	96	-	5.01	6.47	0.27	-	-	-	2.59	2.06	4.97	5.64	-	-
Saccharomyces cerevisiae (other)	4	-	0.39	-	-	0.44	0.69	-	-	0.37	-	-	-	-
	24	-	-	0.77	0.76	-	1.28	2.49	0.76	0.78	0.68	0.76	0.73	-
	48	-	-	0.79	-	-	7.14	7.25	2.42	-	-	-	-	-
	72	-	-	-	-	-	-	-	1.76	-	0.66	0.72	-	2.12
	96	-	0.36	-	0.27	-	-	0.29	0.43	-	-	-	-	-
Pichia kudriavzevii	4	-	-	-	-	-	1.04	-	-	-	-	-	-	-
	24	-	-	-	-	-	0.64	-	-	-	-	-	-	-
	48	-	-	-	-	-	-	-	-	-	-	-	-	-
	72	-	-	-	-	-	-	-	-	-	-	-	-	-
	96	-	-	-	-	2.22	-	-	0.43	1.29	2.26	-	0.28	-

The relative populations sizes of the different yeast species were calculated based on a representative number of yeast isolates taken per time point (see Materials and Methods). Multiplication with the total yeast population size led to the population sizes, displayed as log(CFU g^{-1}). Interdelta analysis was used to differentiate between the inoculated and other *Saccharomyces cerevisiae* strains.