Supplementary Text

Fermentation innovation through complex hybridization of wild and domesticated yeasts

Langdon et al.

Genomic Contributions and Isolation Environments

In our dataset, lager-like (*S. cerevisiae* (*Scer*) × *S. eubayanus* (*Seub*)) hybrids were the most abundant, encompassing 56 strains. These lager-like hybrids can be classified into the two lager lineages, Frohberg (Group II 2n *S. cerevisiae* × 2n *S. eubayanus*) or Saaz (Group I 1n *S. cerevisiae* × 2n *S. eubayanus*) [1]. The next most abundant were hybrids of *S. eubayanus* × *S. uvarum* (*Suva*) (n=41), followed by *S. cerevisiae* × *S. kudriavzevii* (*Skud*) hybrids (n=15). We found four strains with contributions from *S. cerevisiae* × *S. eubayanus* × *S. uvarum* and one *S. cerevisiae* (46%) × *S. eubayanus* (43%) × *S. kudriavzevii* (11%) hybrid. Finally, we also found five strains that are four-way hybrids of *S. cerevisiae* × *S. eubayanus* × *S. kudriavzevii* × *S. uvarum*. These complex hybrids each have unique genomic contributions, with the exception of one *S. cerevisiae* (50%) × *S. eubayanus* (5%) × *S. uvarum* (45%) strain (WLP351), which appears to be very similar to a strain recently isolated from Norwegian Kveik [2] (Supplementary Figure 1). We found minor differences in genomic content between sequencing efforts of the same strain by multiple groups, highlighting the instability of hybrid genomes. Among hybrid types, including closely related strains that likely descended from the same hybridization event, we observed considerable ploidy variation (e.g. the Frohberg *S. cerevisiae* content varied from 47% to 69%). Such rapidly accumulating variation is likely at least partly responsible for the varied properties of closely related fermentation strains. Our global analysis focused on broader genome-scale patterns and was not designed to detect very small contributions, such as the

known subtelomeric *S. cerevisiae* introgressions in some *S. eubayanus* × *S. uvarum* hybrids [3–5] and other minor contributions [6,7]. Thus, the complexity of some hybrids is even greater than our conservative analyses suggest.

We found that lager-like strains (*S. cerevisiae* × *S. eubayanus*) were positively associated with beer (p=2.73E-23, χ^2= 106.11). In contrast, *S. eubayanus* × *S. uvarum* hybrids were not significantly associated with beer, wine, or fruits (the most common isolation origins in our study) (p=1, χ^2=1.56). Thus, these two hybrid types highlight considerable differences in isolation origins and products of all hybrids; whereas the *S. cerevisiae* × *S. eubayanus* hybrids have become almost solely associated with brewing, the *S. eubayanus* × *S. uvarum* hybrids are often viewed as brewing contaminants and have a broader niche (Figure 1b).

Translocation Detection

To complement our analysis of translocations detected with paired-end reads, we also constructed de novo genome assemblies for these hybrids and looked for scaffolds where reads from multiple parents assembled. The detection of translocations with paired-end reads likely overestimates the number of translocations because, even with filtering for mapping quality and uniquely mapped reads, repetitive regions or regions conserved between species can drive cross-mapping. However, looking for scaffolds that assembled from multiple species likely underestimates the number of translocations, because many translocation breaks cannot be assembled across due to the presence of large repeats. However, even with these differences, we observed a similar trend; *S. cerevisiae* × *S. eubayanus* × *S. uvarum* and *S. eubayanus* × *S. uvarum* hybrids and introgressed strains were enriched for translocations (χ^2 = 1250.1, p_adj =2.52 E-05), and the most frequent pair of species between which translocations occurred was *S.*

eubayanus and *S. uvarum* (χ^2 = 38.73, p_adj =8.073 E-07). Overall, these two lines of evidence support the conclusion that translocations are more common between *S. eubayanus* and *S. uvarum* sub-genomes, including the complex hybrids with three or four parent species contributions that include *S. eubayanus* and *S. uvarum* sub-genomes.

Population Genomics:

We caution that building a phylogenetic tree with SNPs concatenated from across the genome in a partially outbred system with considerable admixture, reticulate evolution, gene conversion, and loss-of-heterozygosity, such as yeasts, forces network-like relationships into a bifurcating consensus tree. PCA approaches are model-free and better at capturing and representing complex and conflicting evolutionary histories, but they lack a robust statistical framework. Using these approaches in tandem allowed us to visualize and address any conflicting signals these biases produce.

For our population genomic analyses, we used a set of strains that recapitulated the known *S. kudriavzevii* population structure, including a representative of both Asia B and Asia A and several strains from Europe, including nine genomes newly sequenced for this study. In an extended analysis that included all hybrids, we found that the hybrids split into two groups that matched the previously described hybrid groups A and C (Supplementary Figures 3 & 4, Supplementary Files 2 & 4), which had been inferred by multi-locus data to have arisen out of multiple hybridization events [7,8].

The *S. eubayanus* hybrids with major (>50%) contributions all grouped with the Holarctic lineage (Figure 2b, Supplementary Figure 5, Supplementary Files 2 & 5), a clade also recovered when hybrids with minor contributions were included (Supplementary Figure 6,

Supplementary Files 2 & 6). The connection between the Holarctic lineage and lager strains had been seen on a smaller scale with a few lager strains [9–11]. The number of unique hybrids with *S. eubayanus* contributions indicates that there must have been multiple hybridization and backcrossing events, including those leading to the lager-brewing lineages, backcrossing with *S. uvarum*, and at least one event with a *S. cerevisiae* × *S. kudriavzevii* hybrid. Despite the diversity in chromosome composition, we found that nucleotide diversity in the *S. eubayanus* sub-genome of these hybrids was lower (adjusted π= 0.0208) than the wild populations of *S. eubayanus* (adjusted π =~0.35) and domesticated lineages of *S. cerevisiae* (e.g. Wine yeasts [12], adjusted π = 0.0928). Among the hybrids, the complex hybrids have retained the most nucleotide diversity (adjusted π = 0.0757). Between the lager-brewing lineages, we found that the Saaz lineage had higher nucleotide diversity (adjusted π= 0.0115) than the Frohberg lineage (adjusted π= 0.0028).

Using this expanded collection of interspecies hybrids, we confirmed and extended the finding that all industrially relevant hybrids for *S. uvarum* have arisen out of the *S. uvarum* Holarctic lineage [13]. In contrast to the distinct clustering we observed for the *S. eubayanus* sub-genomes of many of these same hybrids, the *S. uvarum* sub-genomes from these hybrids did not all strongly cluster together.

The barrier to backcrossing with *S. uvarum* after hybridization is less strict than other species barriers in the genus, while the commonality of *S. uvarum* in the Northern Hemisphere provided ample opportunities for *S. uvarum* hybrids to integrate wild diversity. In contrast, other interspecies hybrids seem not to have been able to do so because of *S. kudriavzevii*'s limited genetic relatedness to other *Saccharomyces* and because of *S. eubayanus*' apparent absence or rarity in Europe. Most *S. kudriavzevii* and *S. eubayanus* interspecies hybrids are also allopolyploids, which often rapidly evolve aneuploidy and low spore viability, whereas most *S.*

uvarum interspecies hybrids seem to be relatively fertile homoploid descendants of crosses between primarily *S. eubayanus* and *S. uvarum*. The complexity of these *S. uvarum* hybrids provides an excellent example of how, given opportunity and genetic relatedness, interspecies hybridization will occur and thrive in industrial and other stressful environments.

S. cerevisiae Relationships

We observed two groups of *S. cerevisiae* × *S. kudriavzevii,* correlating with their origin of isolation and memberships in groups A and C [7]. The *S. cerevisiae* sub-genomes of the complex three or four parent species hybrids isolated from wine (S6U) and cider (CID1 and CBS2834) also grouped with the Wine lineage, while the complex hybrids from beer (n=5), the wild (NRRL Y-646), and fruit (NRRL Y-846) grouped with the Beer2 lineage. Specifically, the pattern of genomic completeness in the four-species hybrids suggests that hybrids of *S. cerevisiae* × *S. kudriavzevii* and hybrids of *S. eubayanus* × *S. uvarum* underwent a third hybridization event, leading to complex four-species hybrids.

Lager-brewing yeast nucleotide diversities were considerably lower than the other Ale/Beer1 lineages (adjusted π= ~0.25), but the Saaz lineage (adjusted π= 0.0128) retained more nucleotide diversity than the Frohberg lineage (adjusted π= 0.0035). The low nucleotide diversity of the Frohberg lineage could be the result of a more recent bottleneck and the ongoing dominance of this lineage in commercial beer production where trading and sharing desirable strains is common. To investigate the origin of the *S. cerevisiae* contributions to lager-brewing yeasts, we completed a targeted analysis of the Ale/Beer1 group with a Frohberg dataset that was subsampled to balance the Saaz strains (using the full set of Frohberg strains obscured the structure of Ale/Beer1 due to the low Frohberg diversity) (presented in main text).

Pan-Genome Analyses

In our analyses of novel genomic content (content not found in the parent reference genomes), we found that this content was mostly from non-reference *Saccharomyces* strains. As would be expected, we found an enrichment in *S. cerevisiae, S. kudriavzevii*, and *S. eubayanus* non-reference genetic material, but we also identified novel genetic material whose closest hits were *Saccharomyces paradoxus* (χ^2 = 11510, p_val = 2.2E-16). *S. paradoxus* contributions are known in other industrial *S. cerevisiae* lineages [14]. Some of the *S. cerevisiae* parents had small *S. paradoxus* introgressions that fell below our threshold of detection with sppIDer but were nonetheless discoverable with our pan-genome analyses. Additionally, we recovered the known horizontal gene transfers from *Torulaspora microellipsoides* [15–17] in a few of our hybrids.

Maltotriose Utilization

Maltotriose transport into the cell seems to be the main bottleneck to maltotriose utilization, so we investigated the phylogenetic origin of two key maltose/maltotriose transporters known to exist in lager brewing yeasts, *MTT1* and *AGT1* [18–21]. *MTT1* was originally reported to be a lager-specific gene [22], but here we recovered *MTT1* alleles from nearly 100 *S. cerevisiae* strains (Supplementary Figure 6). In particular, *MTT1* alleles from the Belgian/German/Alt-Kolsch strains were >99% similar to the *MTT1* allele found in lager strains, strongly suggesting that the *S. cerevisiae* parent of the lager lineages contributed this crucial gene. All lager strains have retained a *S. eubayanus* Holarctic *AGT1* gene predicted to be functional, while the *S. cerevisiae AGT1* gene has a premature stop codon in all lager strains [23,24]. Even though some *S. eubayanus* strains possess a Holarctic allele of *AGT1*, no wild strains of *S.*

eubayanus are known to utilize maltotriose, and the *S. cerevisiae* genome is required to confer this trait to synthetic hybrids [25–28].

Mitochondrial Genomes

Collectively, the mitochondrial genomes of these hybrid and introgressed strains belong to nineteen non-*S. cerevisiae* haplotypes (Supplementary Figure 14 & Supplementary File 17), and notably, all lager strains inherited the Holarctic *S. eubayanus* mitochondrial genome (Figure 3 & Supplementary Figure 14).

Many hybrids and introgressed strains inherited mitochondrial genomes from a parent who did not contribute the most nuclear content. For comparison, *S. cerevisiae* and *S. uvarum* typically contributed >50% of the total nuclear content and, on average, 85% complete sub-genomes. In contrast, in hybrids with *S. kudriavzevii* and *S. eubayanus* contributions, these species contributed on average ~1/3 of the total genomic content and <2/3 complete sub-genomes.

4-Vinyl Guaiacol (4-VG)

Pad1 produces a flavin-derived cofactor that is required by Fdc1 for the decarboxylation of phenylacrylic acids, including ferulic acid [29]. Therefore, the disruption of either gene prevents the production of 4VG. Previous studies have shown that most ale-brewing yeasts have inactivated these genes with insertions and premature stop codons [30,31], suggesting that this trait is under strong selection in many brewing yeasts. In the development of new brewing yeasts, the elimination of 4VG production is highly desirable [28,32].

Trappist-style beers made with *S. cerevisiae* × *S. kudriavzevii* hybrids are prized for their phenolic flavor [7]. Therefore, we were not surprised to find that, in a subset of the *S. cerevisiae* × *S. kudriavzevii* hybrids, we retrieved predicted functional *FDC1* and *PAD1* alleles similar or identical to haplotypes also present in Wine and Beer2 strains (Figure 4d & e, Supplementary Figure 15 & Supplementary File 17). Interestingly, these strains have completely lost the *S. kudriavzevii* alleles and only retained *S. cerevisiae* alleles predicted to be functional (Figure 4a). For *S. eubayanus* × *S. uvarum* hybrids, we found that these strains inherited *FDC1* and *PAD1* genes that were predicted to be functional from one species or the other (Figure 4b). *S. eubayanus* × *S. uvarum* hybrids are often viewed as contaminants in the brewing process, and their functional *PAD1* and *FDC1* alleles may contribute to undesired flavors in some cases.

In most Frohberg lager strains, we recovered a *S. cerevisiae* haplotype of *PAD1* predicted to be functional and a *S. cerevisiae* allele of *FDC1* predicted to be non-functional due to a frameshift mutation (Figure 4d & e, Supplementary Figure 15 & Supplementary File 17). This specific frameshift mutation is also found in many Ale/Beer1 strains and some Bread/Mixed strains, indicating that Frohberg lager strains inherited a preexisting loss-of-function allele from their *S. cerevisiae* parent. Interestingly, we were not able to recover any *PAD1* or *FDC1* alleles for the Saaz lineage. When we analyzed the coverage over the subtelomeric region of *S. cerevisiae*'s chrIV, where *S. cerevisiae PAD1* and *FDC1* are located, we found that all Saaz and two Frohberg strains had completely lost this region (Figure 4c). When we analyzed the *S. eubayanus* subtelomeric region of chrXIII, where *S. eubayanus PAD1* and *FDC1* reside, we found that this region had been lost in all lager strains (Figure 4c, Supplementary Figure 2). These losses likely occurred as break-induced replication led to translocations with the homologous subtelomeric regions of the *S. cerevisiae* chrXIII, which lack *PAD1* and *FDC1*.

These translocations, which occur in different locations in different strains, were previously observed in a handful of Saaz and Frohberg strains [33–35], but our study is the first to infer any phenotypic significance.

References

1. Dunn, B. & Sherlock, G. Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res.* **18**, 1610–1623 (2008).
2. Krogerus, K., Preiss, R. & Gibson, B. A unique *Saccharomyces cerevisiae* × *Saccharomyces uvarum* hybrid isolated from norwegian farmhouse beer: Characterization and reconstruction. *Front. Microbiol.* **9**, 1–15 (2018).
3. Libkind, D. *et al.* Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14539–44 (2011).
4. Nguyen, H. V. & Boekhout, T. Characterization of *Saccharomyces uvarum* (Beijerinck, 1898) and related hybrids: Assessment of molecular markers that predict the parent and hybrid genomes and a proposal to name yeast hybrids. *FEMS Yeast Res.* **17**, 1–19 (2017).
5. Nguyen, H. V., Legras, J. L., Neuvéglise, C. & Gaillardin, C. Deciphering the hybridisation history leading to the lager lineage based on the mosaic genomes of *Saccharomyces bayanus* strains NBRC1948 and CBS380 T. *PLoS One* **6**, (2011).
6. Borneman, A. R., Forgan, A. H., Kolouchova, R., Fraser, J. A. & Schmidt, S. A. Whole Genome Comparison Reveals High Levels of Inbreeding and Strain Redundancy Across the Spectrum of Commercial Wine Strains of *Saccharomyces cerevisiae*. *G3* **6**, 957–971 (2016).
7. Peris, D., Pérez-Torrado, R., Hittinger, C. T., Barrio, E. & Querol, A. On the origins and

industrial applications of *Saccharomyces cerevisiae* × *Saccharomyces kudriavzevii* hybrids. *Yeast* **35**, 51–69 (2018).

8. Peris, D., Lopes, C. A., Arias, A. & Barrio, E. Reconstruction of the Evolutionary History of Saccharomyces cerevisiae x S. kudriavzevii Hybrids Based on Multilocus Sequence Analysis. *PLoS One* **7**, (2012).
9. Peris, D. *et al.* Population structure and reticulate evolution of *Saccharomyces eubayanus* and its lager-brewing hybrids. *Mol. Ecol.* **23**, 2031–2045 (2014).
10. Peris, D. *et al.* Complex Ancestries of Lager-Brewing Hybrids Were Shaped by Standing Variation in the Wild Yeast *Saccharomyces eubayanus*. *PLoS Genet.* **12**, (2016).
11. Bing, J., Han, P.-J., Liu, W.-Q., Wang, Q.-M. & Bai, F.-Y. Evidence for a Far East Asian origin of lager beer yeast. *Curr. Biol.* **24**, R380-1 (2014).
12. Almeida, P. *et al.* A population genomics insight into the Mediterranean origins of wine yeast domestication. *Mol. Ecol.* **24**, 5412–5427 (2015).
13. Almeida, P. *et al.* A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum*. *Nat. Commun.* **5**, 4044 (2014).
14. Peter, J. *et al.* Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* **556**, 339–344 (2018).
15. Novo, M. *et al.* Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proc. Natl. Acad. Sci.* **106**, 16333–16338 (2009).
16. Marsit, S. & Dequin, S. Diversity and adaptive evolution of *Saccharomyces* wine yeast: a review. *FEMS Yeast Res.* **15**, 1–12 (2015).
17. Legras, J. L. *et al.* Adaptation of *S. cerevisiae* to fermented food environments reveals

remarkable genome plasticity and the footprints of domestication. *Mol. Biol. Evol.* **35**, 1712–1727 (2018).

18. Han, E.-K., Cotty, F., Sottas, C., Jiang, H. & Michels, C. A. Characterization of AGT1 encoding a general alpha-glucoside transporter from *Saccharomyces*. *Mol. Microbiol.* **17**, 1093–1107 (1995).

19. Salema-Oom, M., Pinto, V. V., Gonçalves, P. & Spencer-Martins, I. Maltotriose Utilization by Industrial. *Society* **71**, 5044–5049 (2005).

20. Baker, E. *et al.* The genome sequence of *Saccharomyces eubayanus* and the domestication of lager-brewing yeasts. *Mol. Biol. Evol.* **32**, 2818–2831 (2015).

21. Brown, C. A., Murray, A. W. & Verstrepen, K. J. Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts. *Curr. Biol.* **20**, 895–903 (2010).

22. Dietvorst, J., Londesborough, J. & Steensma, H. Y. Maltotriose utilization in lager yeast strains: MTTI encodes a maltotriose transporter. *Yeast* **22**, 775–788 (2005).

23. Vidgren, V., Ruohonen, L. & Londesborough, J. Characterization and functional analysis of the MAL and MPH loci for maltose utilization in some ale and lager yeast strains. *Appl. Environ. Microbiol.* **71**, 7846–7857 (2005).

24. Vidgren, V., Viljanen, K., Mattinen, L., Rautio, J. & Londesborough, J. Three Agt1 transporters from brewer's yeasts exhibit different temperature dependencies for maltose transport over the range of brewery temperatures (0-20 °C). *FEMS Yeast Res.* **14**, 601–613 (2014).

25. Hebly, M. *et al. S. cerevisiae* × *S. eubayanus* interspecific hybrid, the best of both worlds and beyond. *FEMS Yeast Res.* **15**, 1–14 (2015).

26. Brickwedde, A. *et al.* Evolutionary Engineering in Chemostat Cultures for Improved

Maltotriose Fermentation Kinetics in Saccharomyces pastorianus Lager Brewing Yeast. *Front. Microbiol.* **8**, 1–15 (2017).

27. Krogerus, K., Magalhães, F., Vidgren, V. & Gibson, B. New lager yeast strains generated by interspecific hybridization. *J. Ind. Microbiol. Biotechnol.* **42**, 769–778 (2015).

28. Krogerus, K., Magalhães, F., Vidgren, V. & Gibson, B. Novel brewing yeast hybrids: creation and application. *Appl. Microbiol. Biotechnol.* **101**, 65–78 (2017).

29. Lin, F., Ferguson, K. L., Boyer, D. R., Lin, X. N. & Marsh, E. N. G. Isofunctional Enzymes PAD1 and UbiX Catalyze Formation of a Novel Cofactor Required by Ferulic Acid Decarboxylase and 4-Hydroxy-3-polyprenylbenzoic Acid Decarboxylase. *ACS Chem. Biol.* **10**, 1137–1144 (2015).

30. Gallone, B. *et al.* Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts. *Cell* **166**, 1397-1410.e16 (2016).

31. Gonçalves, M. *et al.* Distinct Domestication Trajectories in Top- Fermenting Beer Yeasts and Wine Yeasts. *Curr. Biol.* **26**, 1–12 (2016).

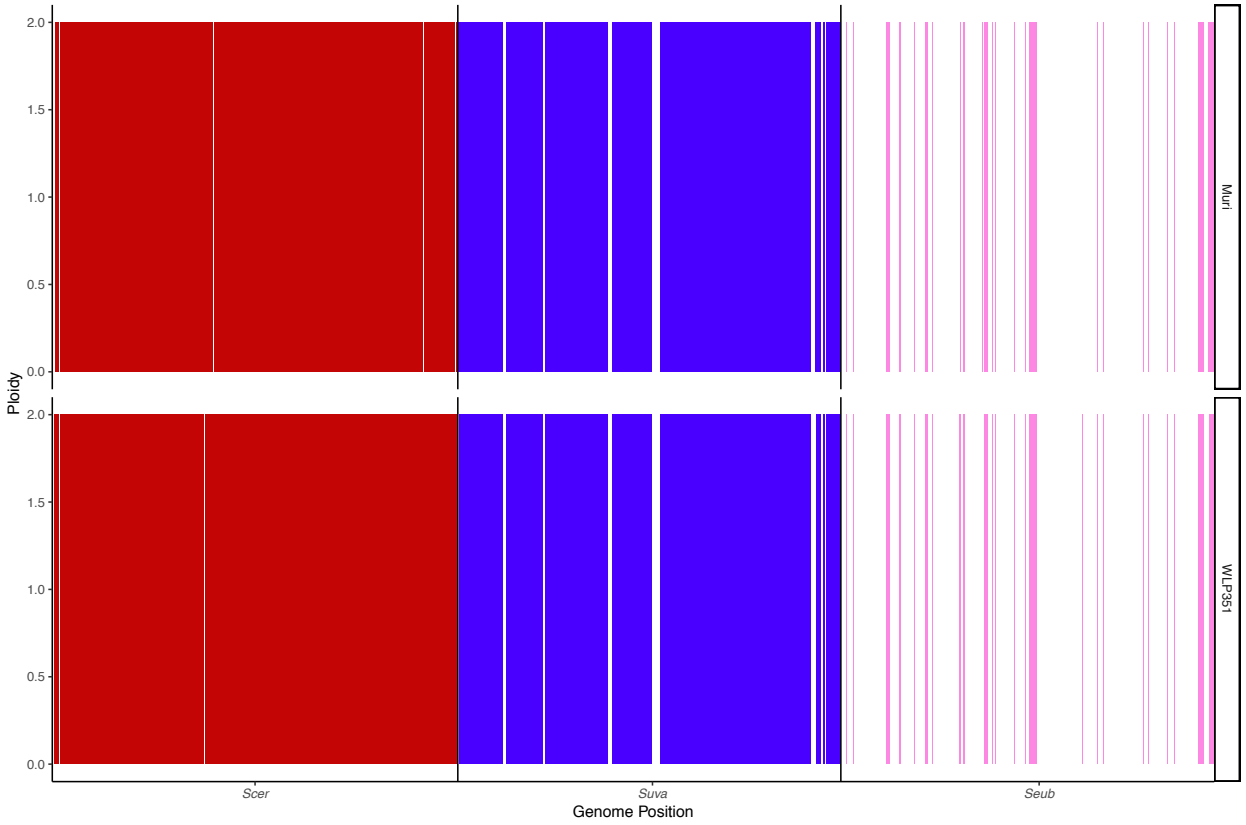
32. Diderich, J. A., Weening, S. M., van den Broek, M., Pronk, J. T. & Daran, J.-M. G. Selection of Pof-*Saccharomyces eubayanus* Variants for the Construction of *S. cerevisiae* × *S. eubayanus* Hybrids With Reduced 4-Vinyl Guaiacol Formation. *Front. Microbiol.* **9**, 1640 (2018).

33. Nakao, Y. *et al.* Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res.* **16**, 115–129 (2009).

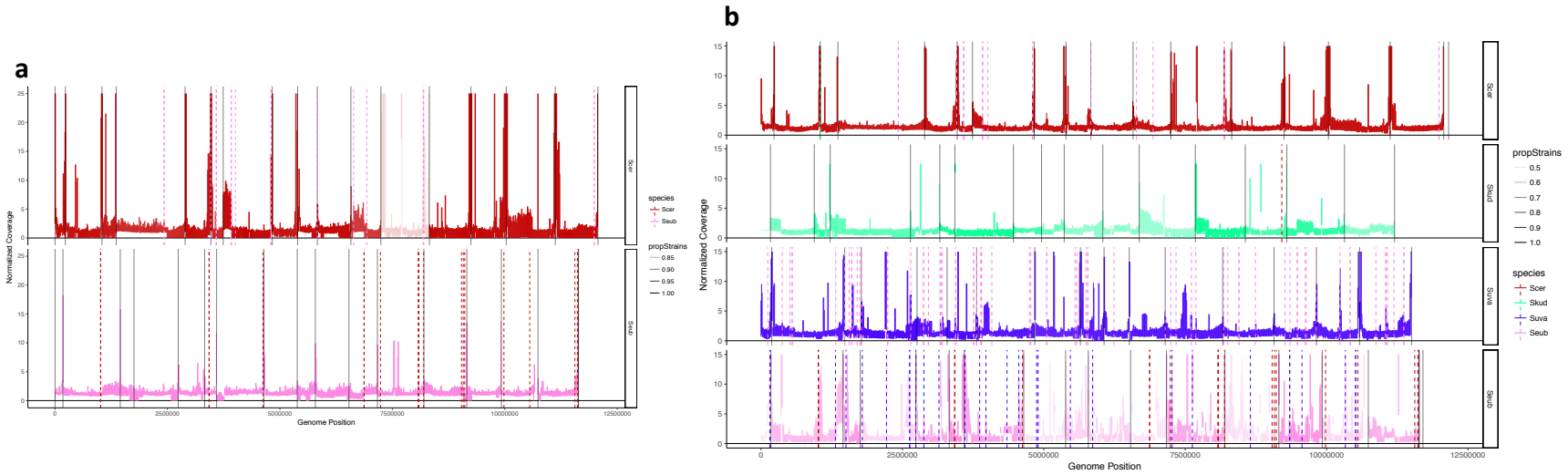
34. Hewitt, S. K., Donaldson, I. J., Lovell, S. C. & Delneri, D. Sequencing and characterisation of rearrangements in three *S. pastorianus* strains reveals the presence of chimeric genes and gives evidence of breakpoint reuse. *PLoS One* **9**, e92203 (2014).

35. Okuno, M. *et al.* Next-generation sequencing analysis of lager brewing yeast strains reveals the evolutionary history of interspecies hybridization. *DNA Res.* **1**, 1–14 (2016).

Supplementary Figures

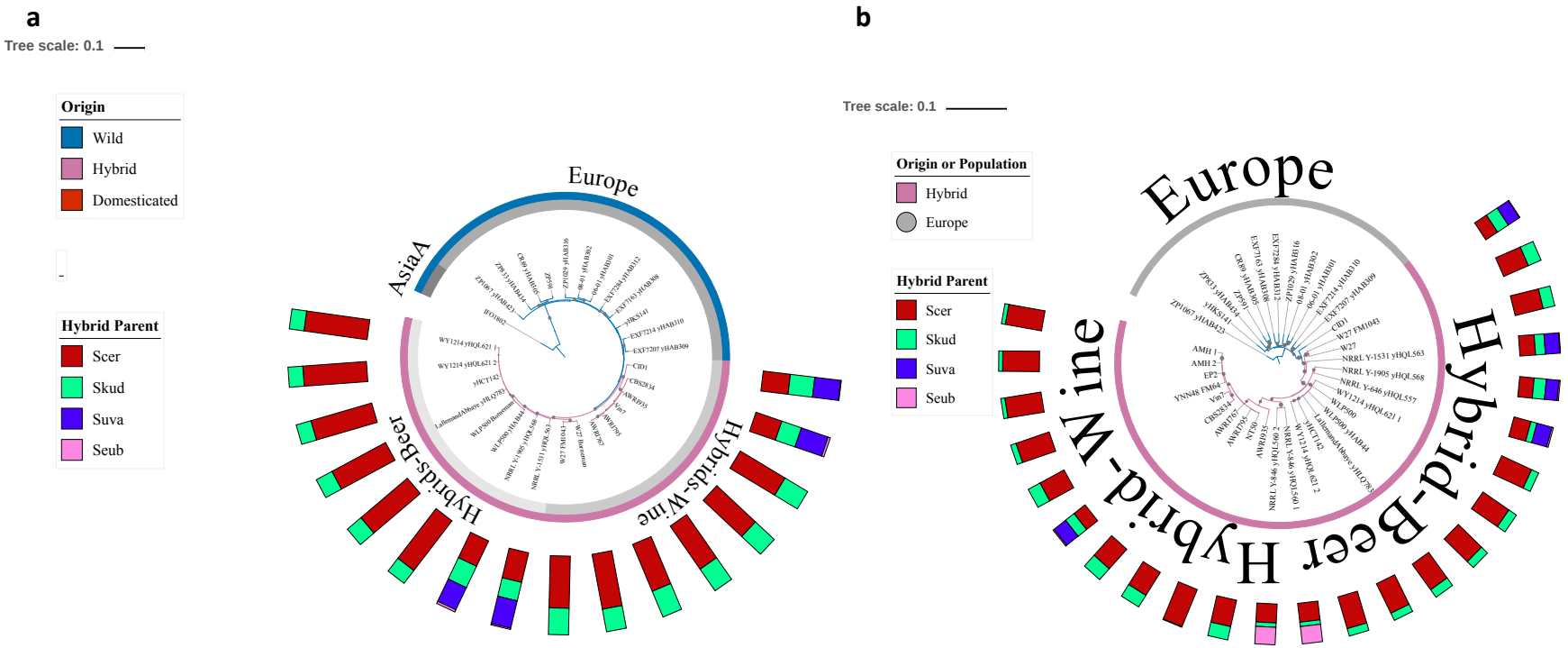


Supplementary Figure 1. Genomic contribution comparison of Muri and WLP351. Modified sppIDer plot, where the y-axis is estimated ploidy, rather than coverage, for the *S. cerevisiae* (50%) × *S. eubayanus* (5%) × *S. uvarum* (45%) strains Muri[61] and WLP351.



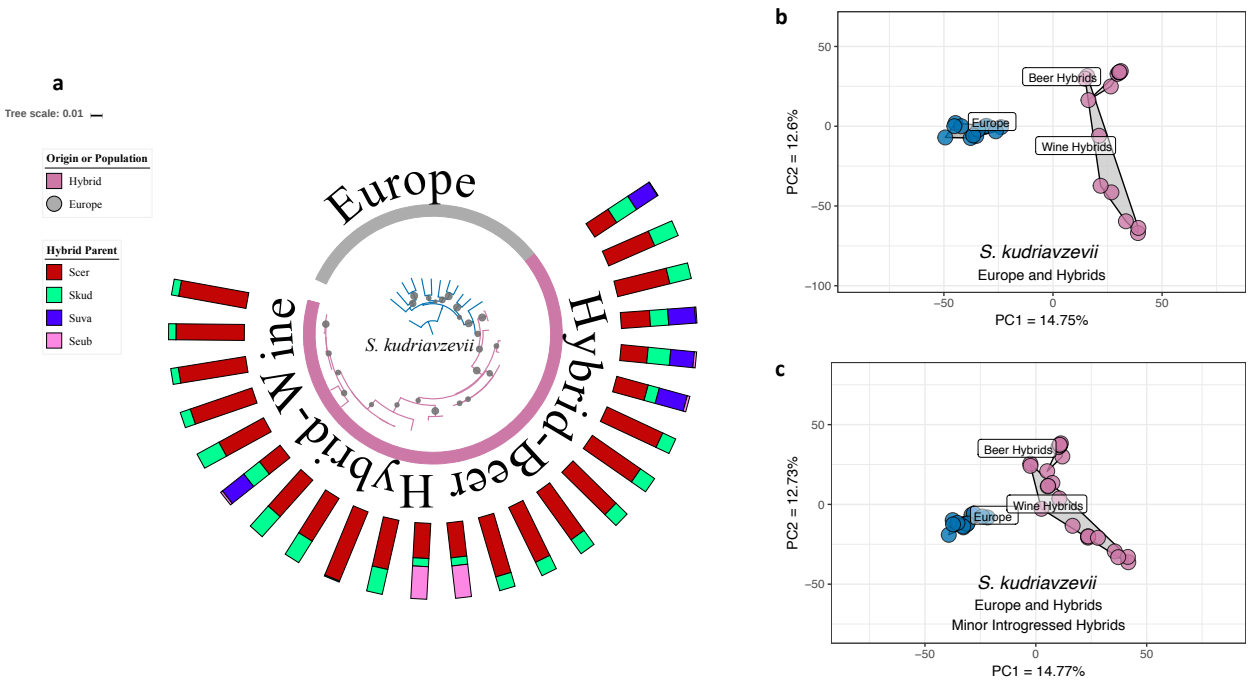
Supplementary Figure 2. Summary of total genomic coverage and shared translocations.

The minimum and maximum normalized coverage of all strains that contain each chromosome are shown as colored bars. Darker chromosomes mean that chromosome is present in more strains. Vertical dotted lines represent translocations that are shared in at least four strains, including between hybrid types. The color of the line represents the reciprocal species. (a) Only lager strains and translocations found only in lagers. (b) All 122 hybrids and interspecies translocations.



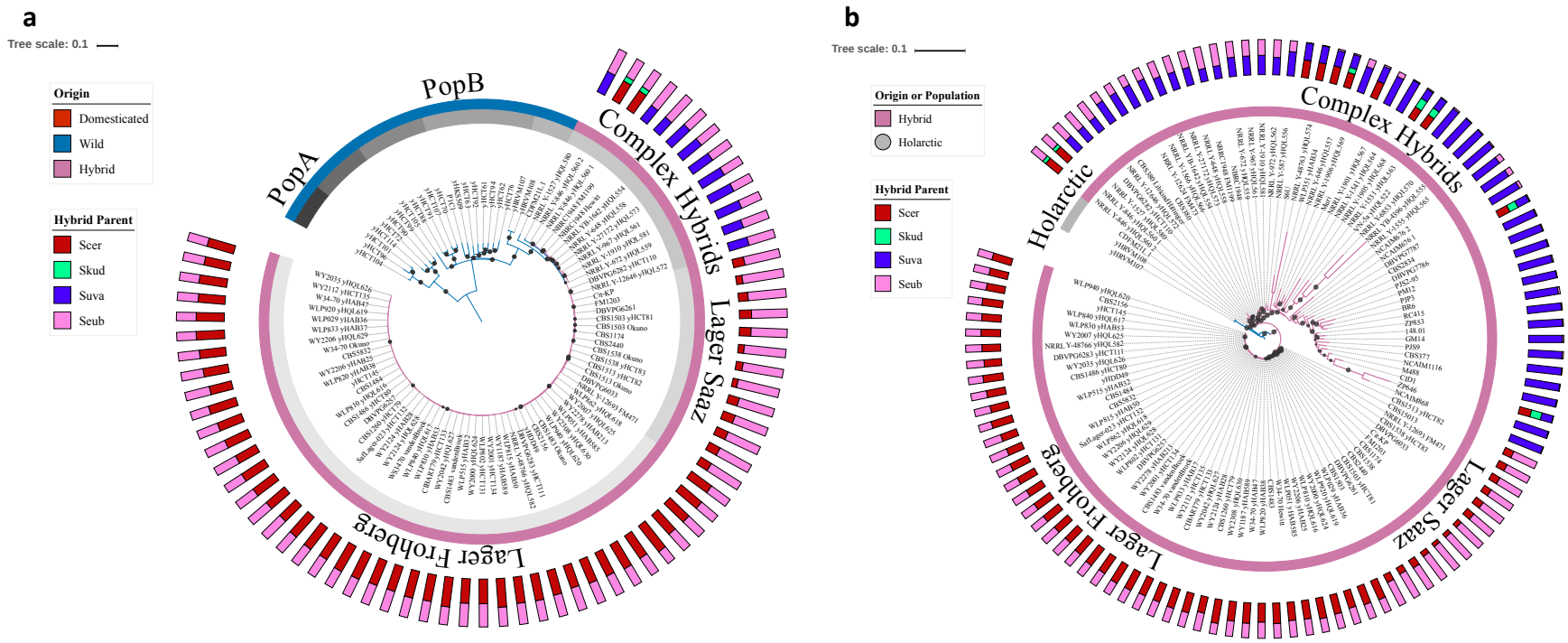
Supplementary Figure 3. Phylogenomic trees for *S. kudriavzevii* with strains labeled.

(a) Phylogeny identical to Figure 2a with strains labeled. (b) Phylogeny identical to Supplementary Figure 4 with strains labeled. Newick files are available as Supplementary File 3 & 4.



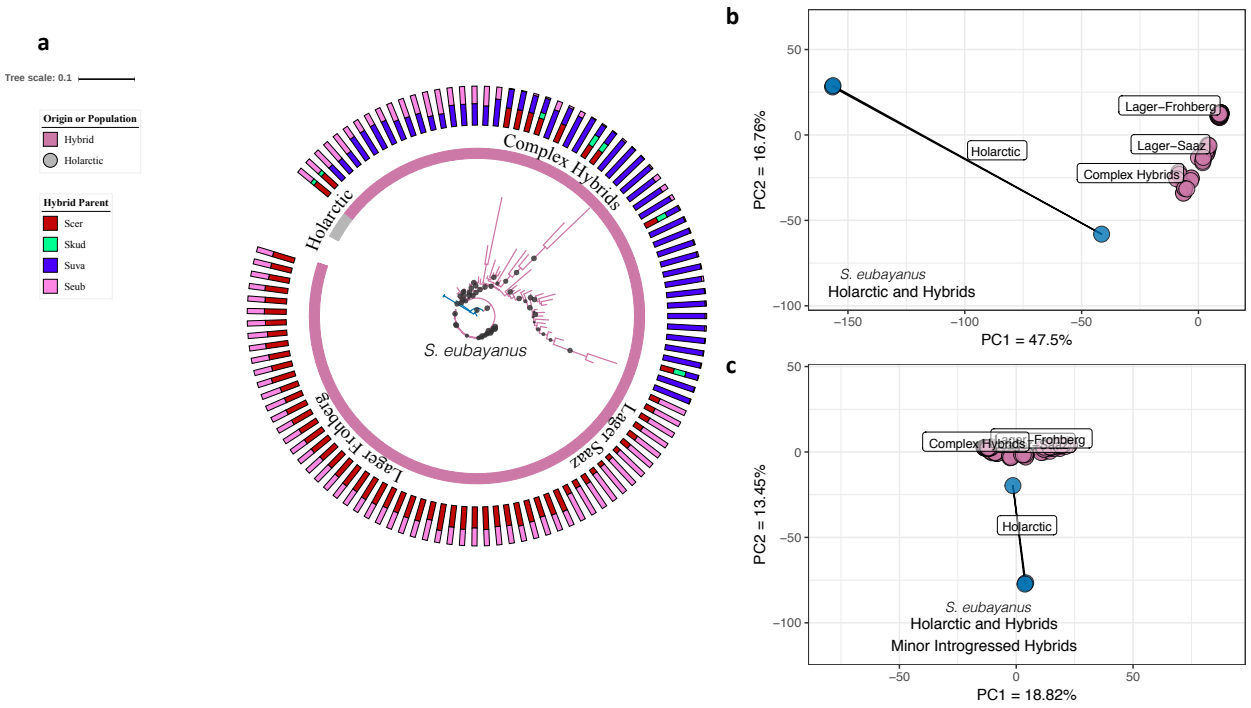
Supplementary Figure 4. Phylogenomic and population placement of hybrids with minor *S. kudriavzevii* contributions.

(a) Phylogenomic tree built with 36 strains and 12,424 SNPs from regions of the genome that exist in at least one minor contributing hybrid. Bootstrap support values >70% are shown as gray dots. Branch colors represent origin of isolation. The inner colors correspond to origin or population. Outer stacked bar plots show the genomic content for each of the hybrids; species colors match Figure 1a. (b) PCA using whole genome data for European *S. kudriavzevii* strains and all major contributor hybrids. (c) PCA using a reduced genome (67%) but including additional minor hybrids. Phylogenies with strain names, Newick formatted files, and data frames used to build PCAs are available as Supplementary Figure 3, Supplementary File 2 & 4.



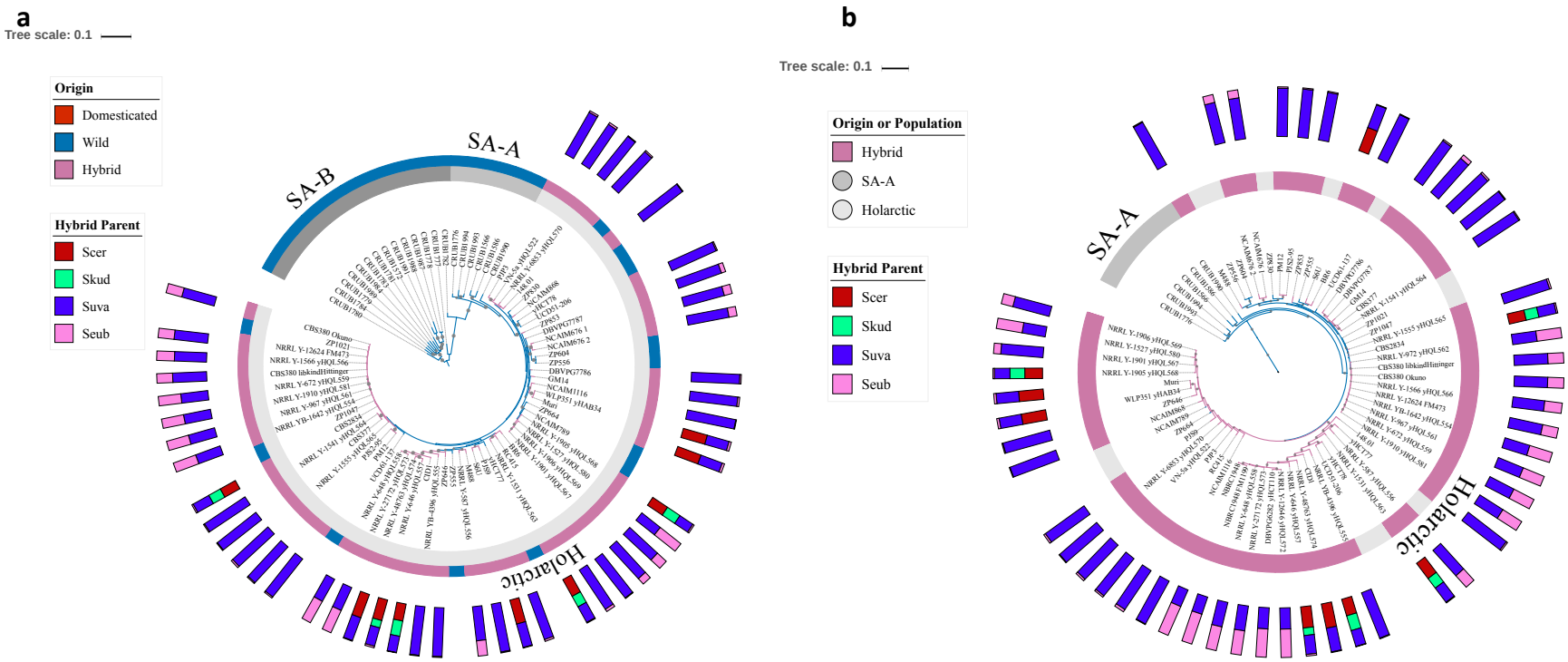
Supplementary Figure 5. Phylogenomic trees for *S. eubayanus* with strains labeled

(a) Phylogeny identical to Figure 2b with strains labeled. (b) Phylogeny identical to Supplementary Figure 6 with strains labeled. Newick files available as Supplementary File 5 & 6.



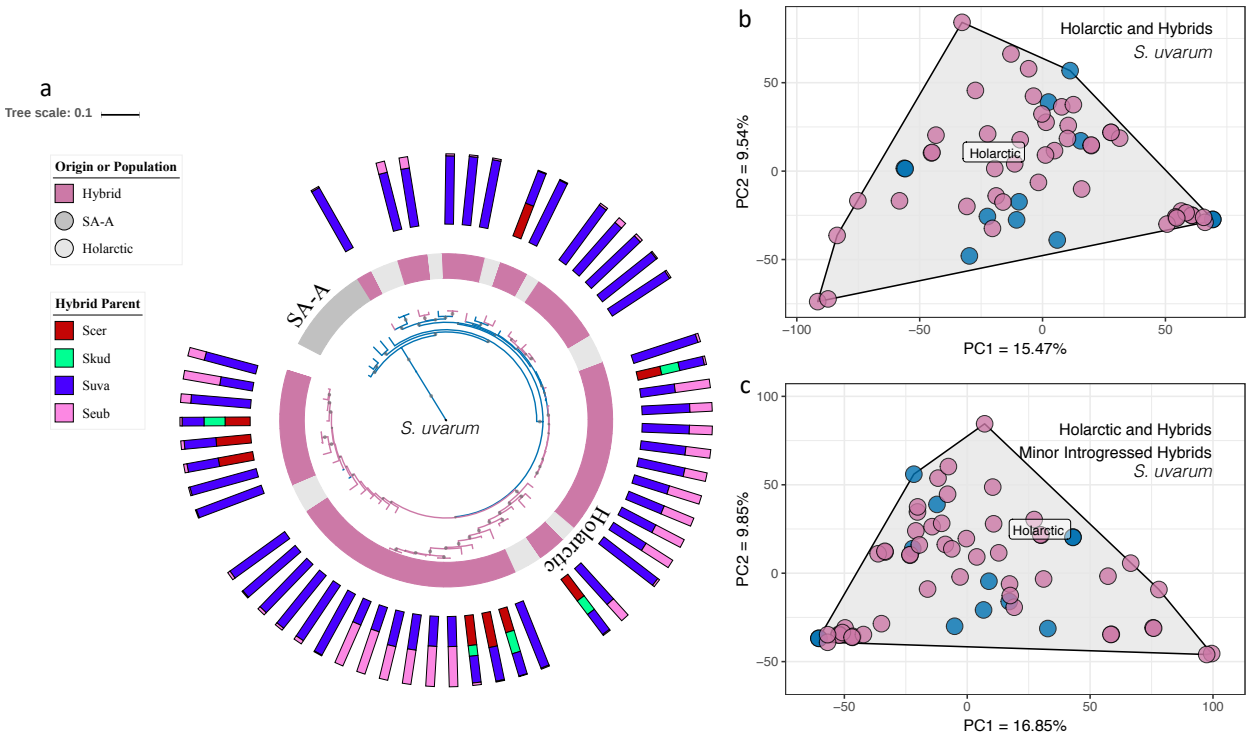
Supplementary Figure 6. Phylogenomic and population placement of hybrids with minor *S. eubayanus* contributions.

(a) Phylogenomic tree built with 112 strains and 69,631 SNPs from regions of the genome that exist in at least one minor contributing hybrid. Bootstrap support values >70% are shown as gray dots. Branch colors represent origin of isolation. The inner colors correspond to origin or population. Outer stacked bar plots show the genomic content for each of the hybrids; species colors match Figure 1a. Long branches are biased by the extensive missing data in hybrids with very small contributions from *S. eubayanus*. (b) PCA using whole genome data for Holarctic *S. eubayanus* strains and all major contributor hybrids. (c) PCA using a reduced genome (25%) but including additional minor hybrids. Phylogenies with strain names, Newick formatted files, and data frames used to build PCAs are available as Supplementary Figure 5, Supplementary File 2 & 6.



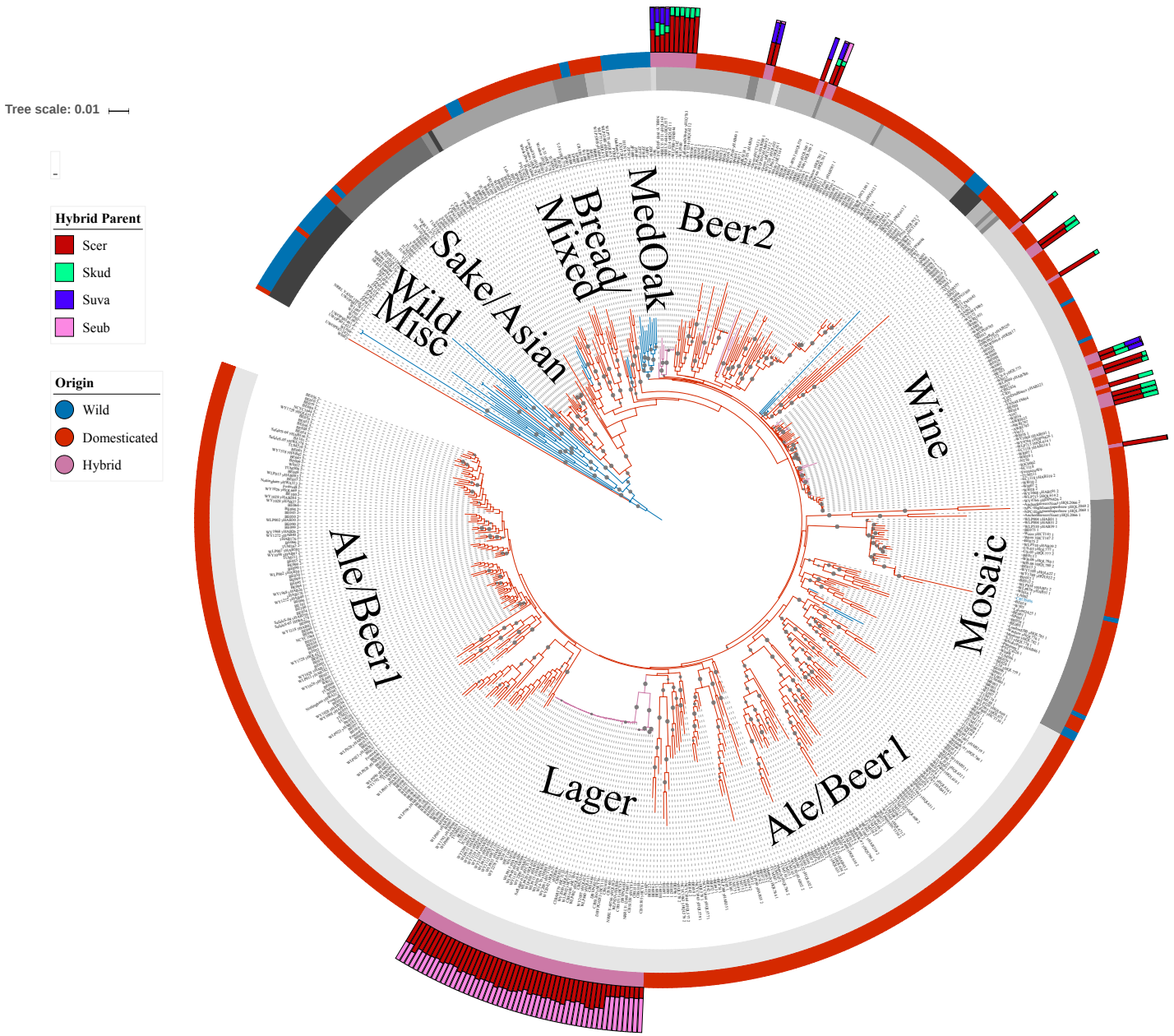
Supplementary Figure 7. Phylogenomic trees for *S. uvarum* with strains labeled.

(a) Phylogeny identical to Figure 2c with strains labeled. (b) Phylogeny identical to Supplementary Figure 8 with strains labeled. Newick files are available as Supplementary File 7 & 8.



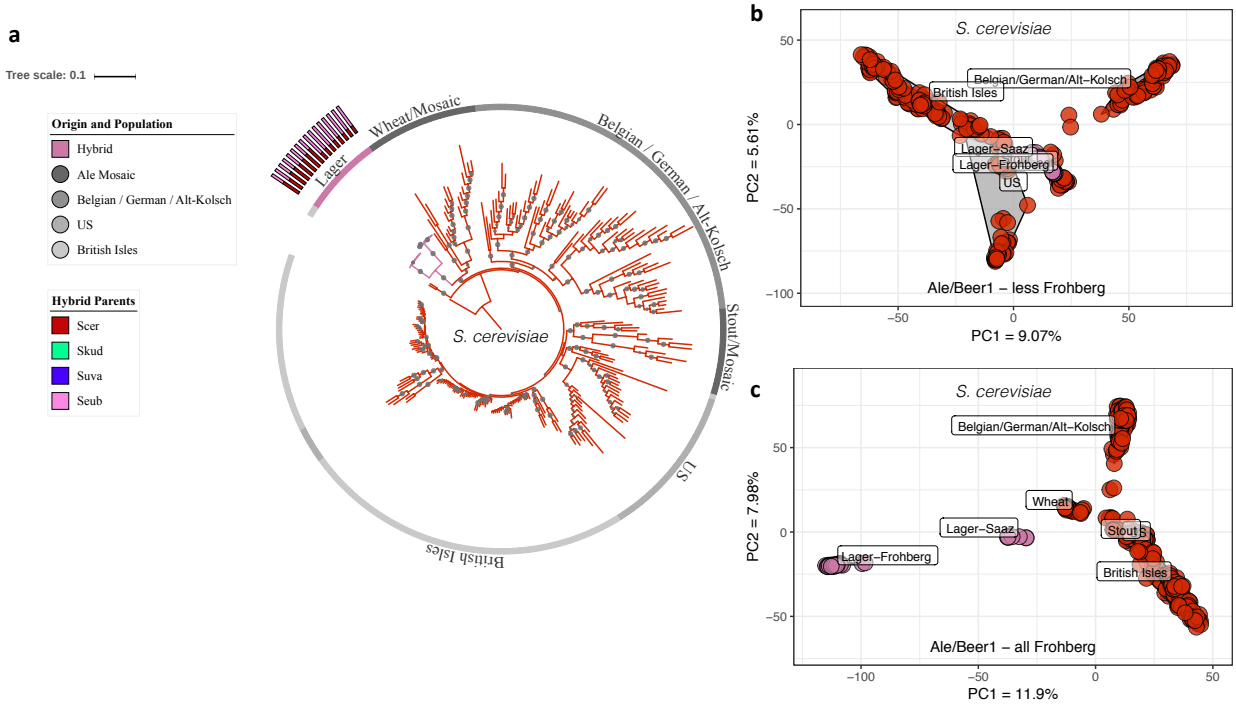
Supplementary Figure 8. Phylogenomic and population placement of hybrids with minor *S. uvarum* contributions.

(a) Phylogenomic tree built with 69 strains and 36,541 SNPs from regions of the genome that exist in at least one minor contributing hybrid. Bootstrap support values >70% are shown as gray dots. Branch colors represent origin of isolation. The inner colors correspond to origin or population. Outer stacked bar plots show the genomic content for each of the hybrids; species colors match Figure 1A. (b) PCA using whole genome data for Holarctic *S. uvarum* strains and all major contributor hybrids. (c) PCA using a reduced genome (84%) but including additional minor hybrids. Phylogenies with strain names, Newick formatted files, and data frames used to build PCAs are available as Supplementary Figure 7, Supplementary File 2 & 8


Tree scale: 0.01
Hybrid Parent
Scer
Skud
Suva
Seub
Origin
Wild
Domesticated
Hybrid
Beer2
MedOak
Bread/
Mixed
Sake/Asian
Wild
Misc
Wine
Mosaic
Ale/Beer1
Lager
Ale/Beer1

Supplementary Figure 9. Phylogenomic tree for full *S. cerevisiae* analysis with strains labeled.

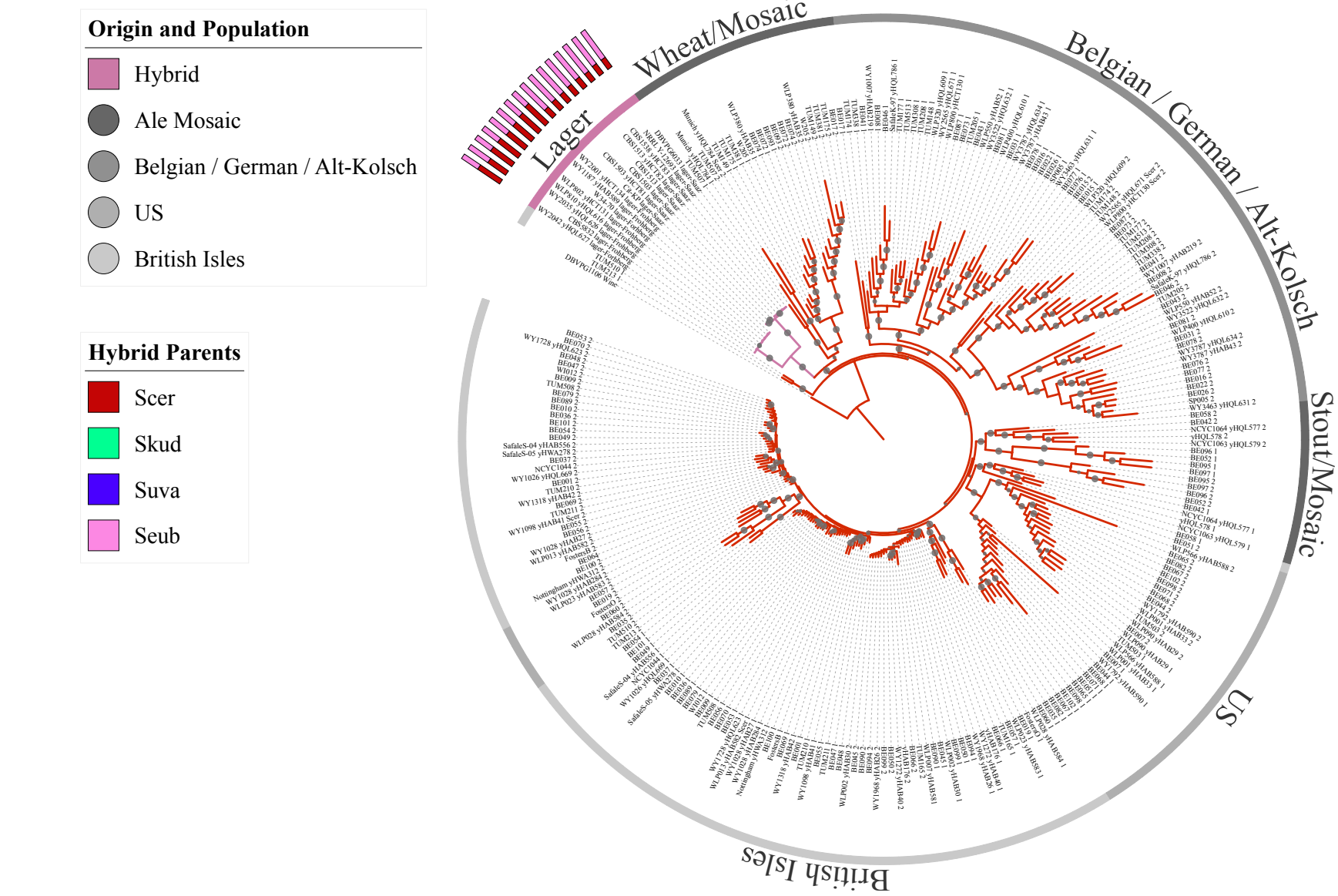
Phylogeny identical to Figure 2d with strains labeled. A Newick file is available as Supplementary File 9.



Supplementary Figure 10. Phylogenomic and population placement of lagers within the Ale/Beer1 clade.

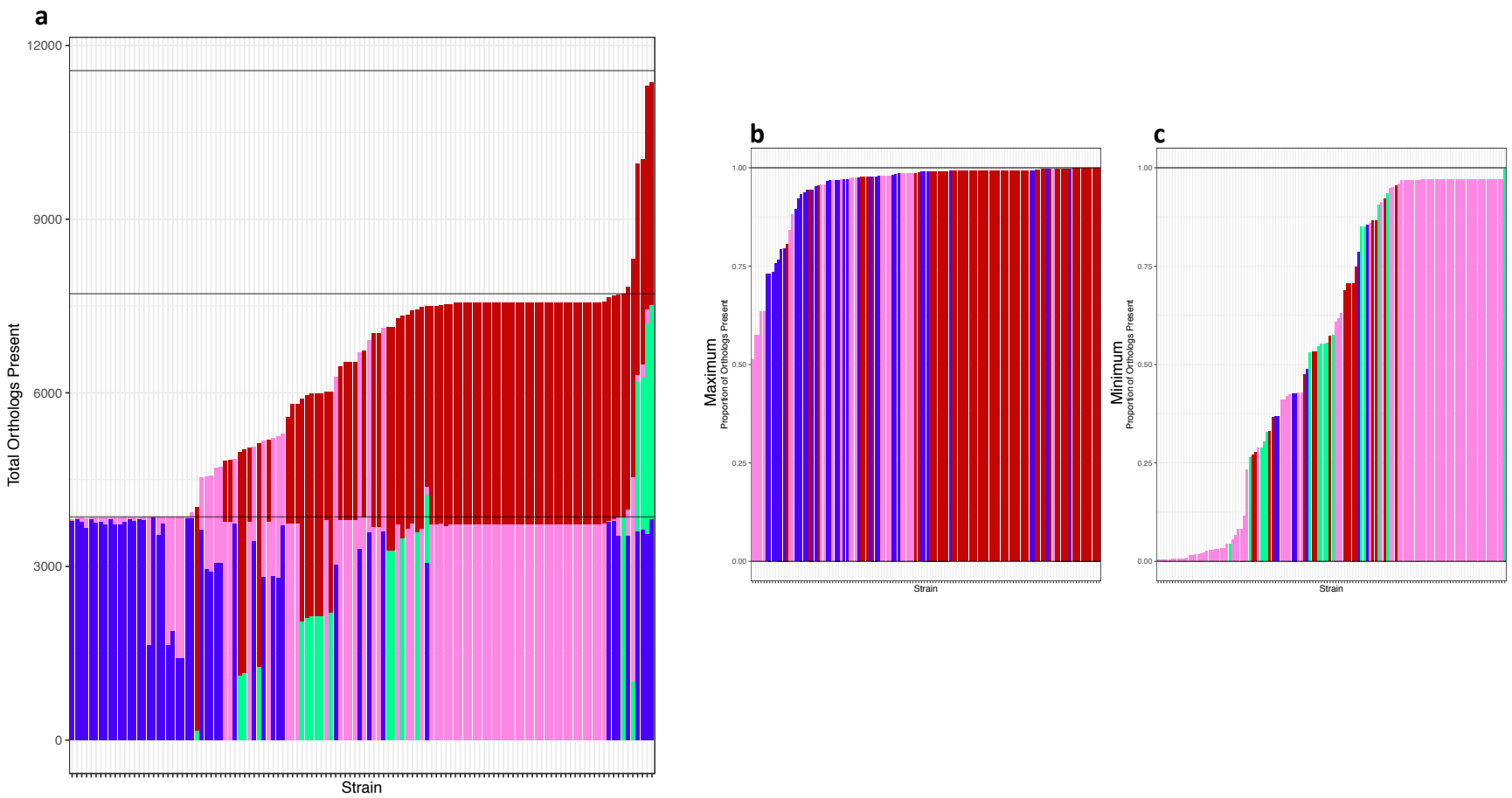
(a) Phylogenomic tree built with 267 strains and 21,953 SNPs from the whole genome. The total number of Frohberg strains was down-sampled to match the same number of Saaz strains. The tree was rooted with the Wine strain DBVPG1106. Bootstrap support values >70% are shown as gray dots. Branch colors represent origin of isolation. The inner colors correspond to origin or population. Outer stacked bar plots show the genomic content for each of the hybrids; species colors match Figure 1a. (b) PCA using whole genome data for Ale/Beer1 strains, all Saaz strains, and the down-sampled set of Frohberg strains. The two lineages of lager strains form separate groups, but they do not cluster with any described geographical lineage of the Ale/Beer1 clade. Pure *S. cerevisiae* Ale/Beer1 strains outside of the labeled lineages are unplaced, including a cluster of Stout strains, Wheat strains, and mosaic strains that our analyses suggest share the

most ancestry with lager-brewing yeasts. (c) PCA using all lager strains. The low diversity in the Frohberg lager strains drives PC1, which led us to balance the dataset by down-sampling this lineage. Phylogenies with strain names, Newick formatted files, and data frames used to build PCAs are available as Supplementary Figure 11, Supplementary File 2 & 10.



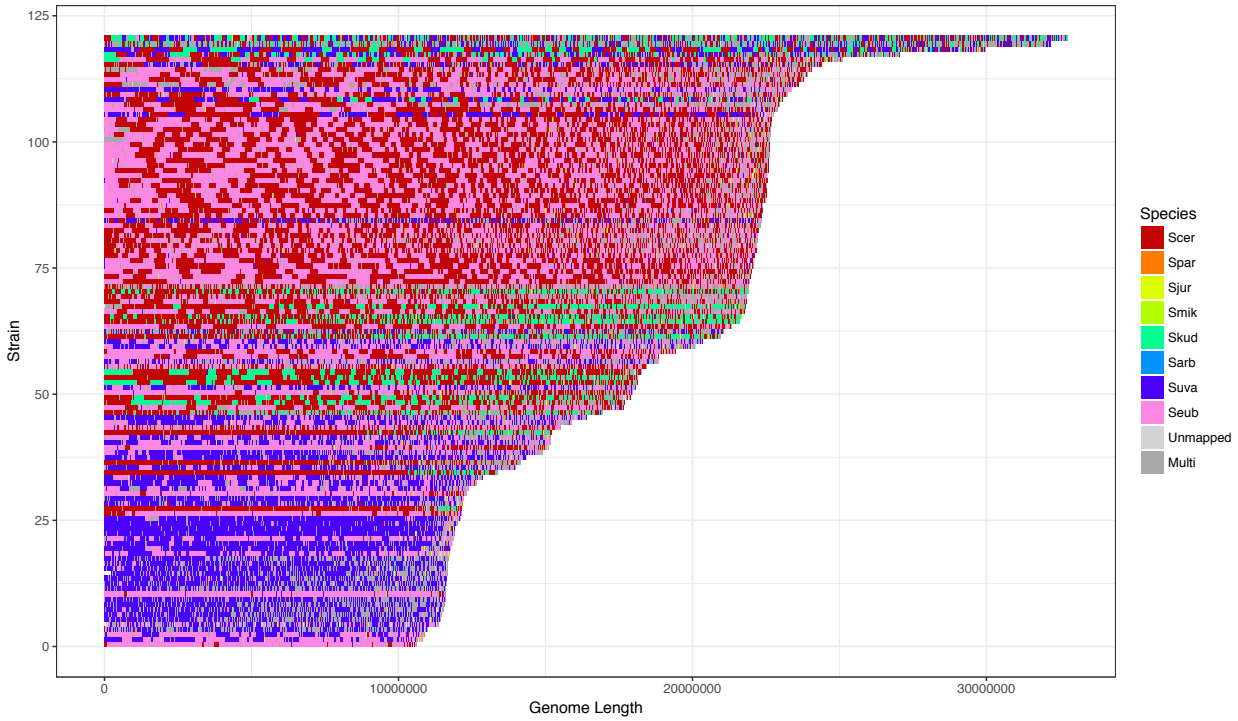
Supplementary Figure 11. Phylogenomic tree for Ale/Beer1 *S. cerevisiae* analysis with strains labeled.

Phylogeny identical to Supplementary Figure 10 with strains labeled. A Newick file is available as Supplementary File 10.



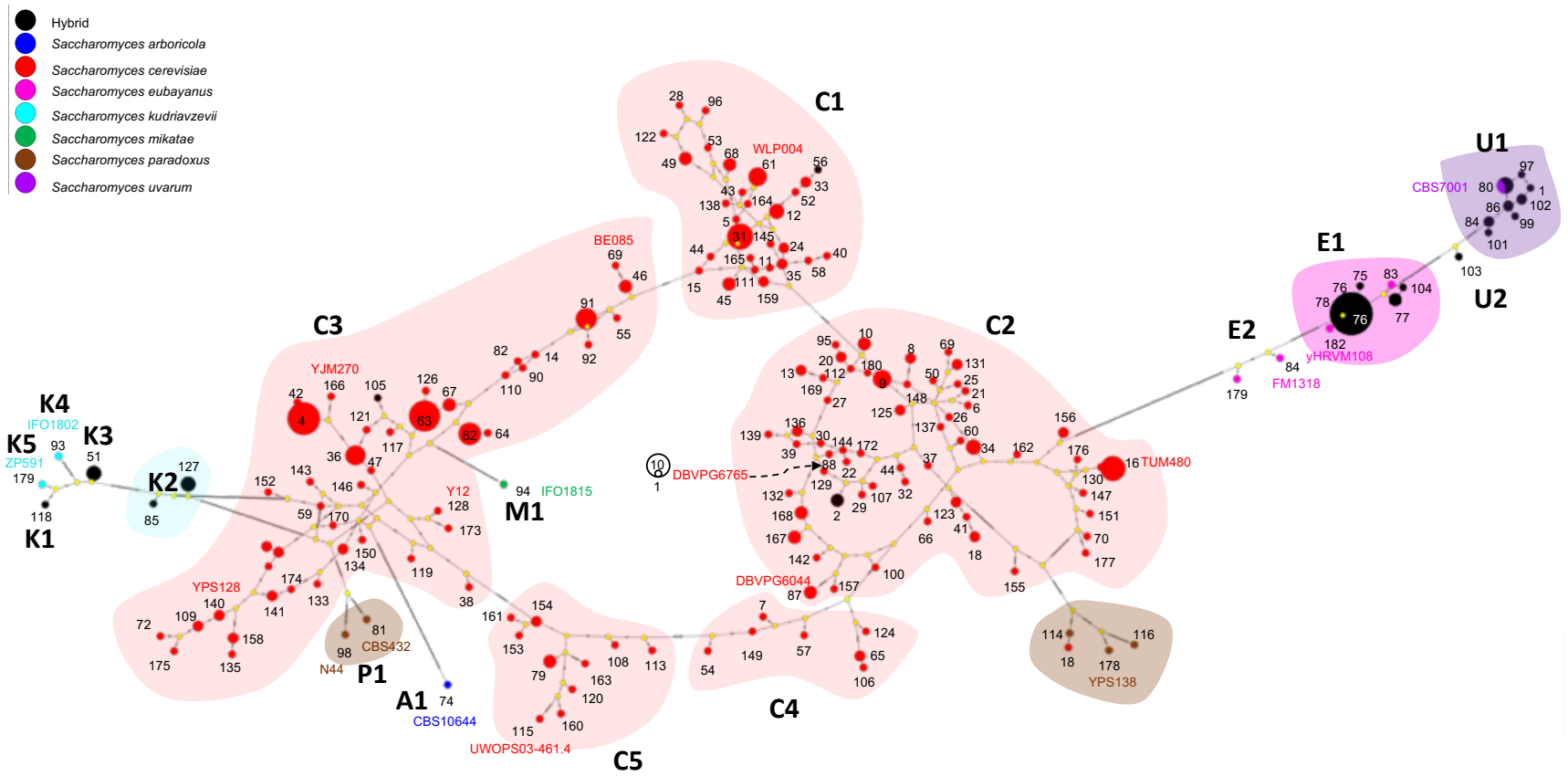
Supplementary Figure 12. 1:1:1:1 orthologs present in hybrid genomes.

(a) Stacked bar chart of all 1:1:1:1 orthologs present in hybrids. Strains are sorted from most to least ortholog content. Completeness of the ortholog set from the species that contributed the most (b) or least (c) orthologs to the strains. Strains are ordered independently in all panel



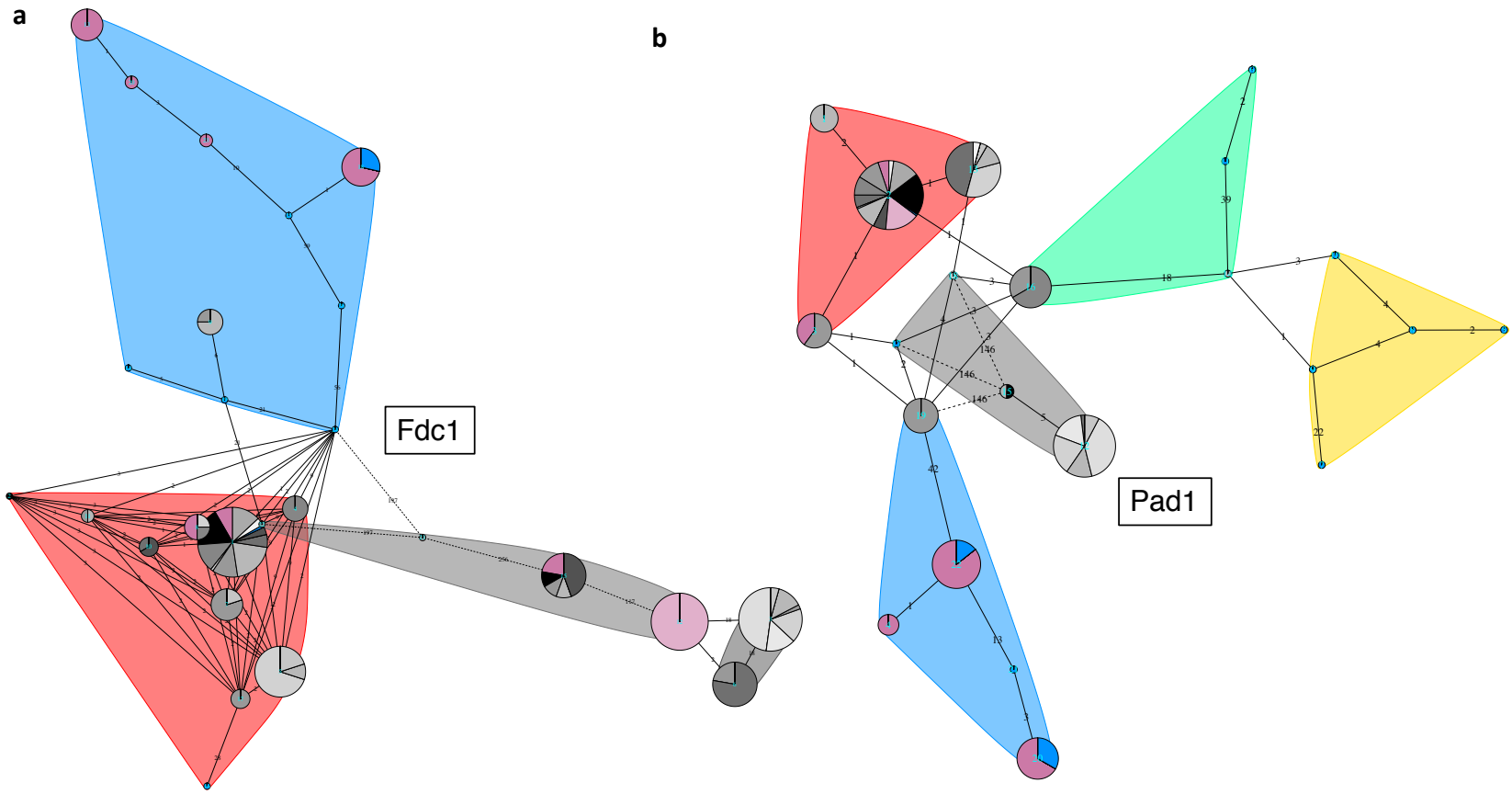
Supplementary Figure 13. Complete de novo genome assembly for all strains.

Total assembled genome for each strain. Regions are colored by which parent could be assigned in the de novo assembly based on the sppIDer results. “Multi” are regions where reads from many species mapped at high coverage. “Unmapped” are novel regions assembled from reads that do not map to parent reference genomes. These “Multi” and “Unmapped” regions constitute small regions interspersed throughout the assembled genomes. For each assembly, contigs are ordered from largest to smallest from left to right.



Supplementary Figure 14. Mitochondrial genome haplotype network.

Six mitochondrial genes were concatenated in 364 wild *Saccharomyces* strains and interspecies hybrids and used to build a TCS[88] phylogenetic network. Haplotype classification is provided in Supplementary File 17. Haplotypes are represented by circles, and circle size is scaled according to the haplotype frequency. Pie charts show the frequency of haplotypes based on species or hybrid designation. The number of mutations separating each haplotype are indicated by lines on the edges connecting the haplotype circles.



Supplementary Figure 15.

Labeled (in turquoise) haplotype networks for *PAD1* and *FDC1*. Edge numbers are the number of amino acid changes. Networks correspond to those used in Figure 4 for the amino acid sequences of (a) Fdc1 and (b) Pad1. (b) A different haplotype network orientation of Figure 4E that increases the visibility of each community and haplotype. Supplementary File 17 contains the key to which strains belong to which haplotype.

Supplementary Files

Supplementary File 1. All hybrids and their parent contributions.

Supplementary File 2. PCA analyses.

Percent explained by each principal component included in column headers.

Supplementary File 3. Newick formatted file of the *S. kudriavzevii* phylogeny with major hybrids.

Supplementary File 4. Newick formatted file of the *S. kudriavzevii* phylogeny with minor hybrids.

Supplementary File 5. Newick formatted file of the *S. eubayanus* phylogeny with major hybrids.

Supplementary File 6. Newick formatted file of the *S. eubayanus* phylogeny with minor hybrids.

Supplementary File 7. Newick formatted file of the *S. uvarum* phylogeny with major hybrids.

Supplementary File 8. Newick formatted file of the *S. uvarum* phylogeny with minor hybrids.

Supplementary File 9. Newick formatted file of the *S. cerevisiae* phylogeny with all strains analyzed.

Supplementary File 10. Newick formatted file of the *S. cerevisiae* phylogeny of just the Ale/Beer1 clade.

Supplementary File 11. Results of Fisher's Exact Test and Bonferroni correction of mitochondrially localized genes.

mtInteracting = nuclear-encoded but mitochondrially localized gene.

Supplementary File 12. Summary of number of 1:1:1:1 orthologs present in each sub-genome.

Supplementary File 13. GO term results of genes found in novel regions of the de novo assembled genomes.

Supplementary File 14. Brewing relevant gene summaries.

"-" Indicates when HybPiper failed to recover and assemble genes for this group or that these assemblies failed our length and coverage cutoffs.

Supplementary File 15. Metadata for all strains newly sequenced in this study.

The "New hybrid" column denotes hybrid genome sequences that are newly published in this study.

Scer = S. cerevisiae, *Spar = Saccharomyces paradoxus*, *Smik = Saccharomyces mikatae*, *Skud = S. kudriavzevii*, *Suva = S. uvarum*, and *Seub = S. eubayanus*.

Supplementary File 16. Published data accession information.

Supplementary File 17. Haplotype key for mitochondrial genomes, *PAD1*, and *FDC1*.

Dataset A only includes strains where *15S rRNA* could be assembled, while Dataset B has *15S rRNA* removed.

Supplementary File 18. Regions used for minor contribution analyses.