Interspecific hybridisation facilitates niche adaptation in beer yeast

Brigida Gallone[1,2,3,4,5*], Jan Steensels[1,2,3*], Stijn Mertens[1,2,3], Maria C. Dzialo[1,2,3], Jonathan L. Gordon[1,2,3], Ruben Wauters[1,2,3], Florian A. Theßeling[1,2,3], Francesca Bellinazzo[1,2,3], Veerle Saels[1,2,3], Beatriz Herrera-Malaver[1,2,3], Troels Prahl[6], Christopher White[6], Mathias Hutzler[7], Franz Meußdoerffer[7], Philippe Malcorps[8], Ben Souffriau[8], Luk Daenen[8], Guy Baele[9], Steven Maere[4,5*] and Kevin J. Verstrepen[1,2,3*].

[1] VIB – KU Leuven Center for Microbiology, Gaston Geenslaan 1, B-3001 Leuven, Belgium.

[2] CMPG Laboratory of Genetics and Genomics, Department M2S, KU Leuven, Gaston Geenslaan 1, B-3001 Leuven, Belgium.

[3] Leuven Institute for Beer Research (LIBR), Gaston Geenslaan 1, B-3001 Leuven, Belgium.

[4] Ghent University, Department of Plant Biotechnology and Bioinformatics, Technologiepark 71, 9052 Ghent, Belgium

[5] VIB Center for Plant Systems Biology, Technologiepark 71, 9052 Ghent, Belgium

[6] White Labs Inc., 9495 Candida Street, San Diego, CA 92126, USA

[7] Research Center Weihenstephan for Brewing and Food Quality, TU München, 85354 Freising, Germany

[8] AB-InBev SA/NV, Brouwerijplein 1, B-3000 Leuven, Belgium

[9] Department of Microbiology and Immunology, Rega Institute, KU Leuven – University of Leuven, Herestraat 49, 3000 Leuven, Belgium

* Corresponding authors:

Steven Maere – steven.maere@ugent.vib.be ORCID: 0000-0002-5341-136X

Kevin J. Verstrepen - kevin.verstrepen@kuleuven.vib.be ORCID: 0000-0002-3077-6219

Abstract

Hybridisation between species often leads to inviable or infertile offspring, yet examples of evolutionary successful interspecific hybrids have been reported in all kingdoms of life. However, many questions on the ecological circumstances and evolutionary aftermath of interspecific hybridisation remain unanswered. In this study, we sequenced and phenotyped a large set of interspecific yeast hybrids isolated from the brewing environments to uncover the influence of interspecific hybridisation in yeast adaptation and domestication. Our analyses demonstrate that several hybrids between *Saccharomyces* species originated and diversified in industrial environments by combining key traits of each parental species. Furthermore, post-hybridisation evolution within each hybrid lineage reflects sub-specialisation and adaptation to specific beer styles, a process that was accompanied by extensive chimerisation between subgenomes. Our results reveal how interspecific hybridisation provides an important evolutionary route that allows swift adaptation to novel environments.

Introduction

A canonical view of Darwinian evolution asserts that new species arise during extended periods of gradual evolution and selection in combination with reproductive isolation. However, amendments to Darwin's original theory have been proposed to incorporate other mechanisms for evolution and speciation, including the emergence of new species through interspecific hybridisation[1–3]. Hybridisation provides a way to rapidly combine distinct phenotypic features from established populations that converge in an unfamiliar ecological niche. Sometimes, a hybrid's unique combination of phenotypes can enable it to thrive in the new environment and outcompete its parental species[4–6].

Examples of rapid niche specialisation via interspecific hybridisation are found across all kingdoms of life. A new hybrid lineage of Darwin's finches with competitive beak morphology emerged in only three generations[7]. Hybridisation of the sunflowers *Helianthus annus* and *Heliantus petiolaris* gave rise to three novel species capable of colonizing previously untapped environments[8–10]. The emerging opportunistic fungal pathogen *Candida metapsilosis* arose from a single hybridisation event between two non-pathogenic parental species[11].

The *Saccharomyces* species complex provides numerous examples of interspecific hybridisation. Despite the high sequence divergence between the species within this genus (up to 20%[12]), prezygotic barriers are weak and species can successfully crossbreed[13,14]. The resulting hybrids are typically infertile, yet viable, and can reproduce asexually by budding[14–17]. Wild *Saccharomyces* interspecific hybrids are occasionally encountered[18,19], but the best-known example, *Saccharomyces pastorianus*, was isolated from an industrial environment[20–23].

Humans have historically utilised the capacity of *Saccharomyces* to convert high concentrations of sugars into ethanol and carbon dioxide to produce a wide variety of fermented products[24–26]. This long-term domestication process has resulted in hundreds of different industrial strains, each with characteristics suitable for specific industrial processes[27–36]. Strains used in the production of lager (pilsner) beer generally belong to the species *S. pastorianus*, a hybrid of *S. cerevisiae* and *Saccharomyces eubayanus*[20–23,37,38]. Two archetypical *S. pastorianus* sublineages exist, dubbed 'Frohberg' and 'Saaz', each with its own distinct fermentation properties[39]. Several other industrial interspecific hybrids have been isolated, including *S. cerevisiae* x *Saccharomyces uvarum* x *S. eubayanus* triple hybrids

in wine and cider, and *S. cerevisiae* x *Saccharomyces kudriavzevii* hybrids in ale beers and wine fermentation vessels[40–47].

A comprehensive analysis of the prevalence, molecular details, and the ecological and evolutionary context of interspecific hybridisation in yeasts is lacking. Here, we report how whole-genome sequencing of more than 200 industrial yeasts revealed that a surprisingly large fraction (~25%) proved to be interspecific hybrids derived from four parental species: *S. cerevisiae, S. kudriavzevii, S. eubayanus,* and *S. uvarum*. The ubiquity of these hybrids and the defined industrial environments they were isolated from, make them an excellent model for studying the role of hybridisation in microbial evolution and adaptation. By combining large-scale phenotyping with our knowledge of industrial niches and beer brewing history, we provide evidence that these hybrids originated in industrial environments and are highly niche-specific. Additionally, we characterised the genomic changes that occurred during their domestication and describe the genetic mechanism leading to a key domestication phenotype, namely flavour production. Our results demonstrate that interspecific hybridisation is an evolutionary strategy that allows swift adaptation to novel niches and opens new avenues for the development of superior industrial yeasts.

Results

Origins and diversity of *Saccharomyces* interspecific beer hybrids

For several years we have been collecting and sequencing yeasts isolated from different industrial niches, including beer, wine, bread, sake, chocolate and liquor fermentations. As expected, the majority of isolates were *S. cerevisiae*, the yeast species most commonly associated with the production of fermented foods and beverages[28,48,49]. However, our analysis revealed that a surprisingly large fraction (~25%) of sequenced isolates were interspecific hybrids. While some were isolated from lager beer fermentations, known to be driven by hybrid yeasts, many were collected from other beer niches, such as Trappist beers, spontaneously fermented 'Lambic' beers, and old beer bottles or equipment (**Table S1**). We identified three distinct species compositions across the hybrids investigated: *S. cerevisiae x S. eubayanus* (*S.cer x S.eub*), *S. cerevisiae x S. kudriavzevii* (*S.cer x S.kud*) and *S. eubayanus x S. uvarum* (*S.eub x S.uva*). Phylogenetic trees were built to position the hybrid subgenomes in the context of their parental species (**Figure 1; Figure S1**).

The *S.cer* subgenomes of the *S.cer x S.eub* (*S. pastorianus*) lager hybrids belong to a well-defined monophyletic clade in the Beer 1 lineage, which mainly contains strains isolated from ale beers[28,29,34] (**Figure 1; Figure S1A**). Within Beer 1, the lager subclade is most closely related to Hefeweizen (German wheat beer) strains and the Belgium/Germany lineage, indicating a Western European origin of lager yeasts. The lager clade of Beer 1 is further separated into the archetypical 'Frohberg' and 'Saaz' lineages[22,50,51], both hallmarked by very low nucleotide diversity (π=1.18E-03 and π=3.65E-04, respectively). The *S.eub* subgenomes of the *S.cer x S.eub* hybrids similarly form a monophyletic group within the previously identified Holarctic clade[52] (**Figure 1; Figure S1B)**, with defined 'Frohberg' and 'Saaz' subclades. The Saaz lineage further divides in two subclades, each harbouring one of the two first *S. pastorianus* strains isolated at Carlsberg in the 19th century: Unterhefe nr. 1 (CBS1513) and nr. 2 (CBS1503) (Table S1, **Supplementary Note**). In accordance with previous reports, the *S.eub* progenitor(s) of lager yeasts appear to be most closely related to Tibetan *S.eub* strains, which may have reached Europe through Silk Road trading[23,50,53]. Interestingly, the origin of domesticated barley (another major beer ingredient) in North-western Europe was also traced to Tibet[54]. However, considerable outcrossing and incomplete lineage sorting among Holarctic *S.eub* strains makes determination of the exact geographical origin of the *S.eub* ancestor(s) of lager yeasts difficult[52].

All but one of the *S.cer* subgenomes of the *S.cer x S.kud* hybrids belong to a monophyletic clade, closely related to the industrial yeast clade Beer 2 (**Figure 1** and **Figure S1A**). Similar to lager yeasts, this 'traditional Belgian beer' clade further divides into two subgroups mainly containing hybrids isolated from either Lambic beers ('spontaneous beer fermentations'; ABI1606 and ABI1525) or Trappist beers ('Trappist ales'; BE105, BE108, BE109, BE116 and ABI1620) (**Figure S1A**). One hybrid bread strain (BR005) clusters within this clade, likely reflecting the historical relationship between brewers and bakers[55] (**Figure S1A**). A close association between bread and ale beer strains was previously found among pure *S.cer* strains as well ('Mixed' clade **Figure S1A**)[28]. The only *S.cer x S.kud* strain not belonging to the monophyletic clade is the VIN7 wine strain, whose *S.cer* subgenome clusters within the Wine clade. This indicates that VIN7 originated independently from the ale and bread hybrids. The phylogenetic relationships of the *S.kud* subgenomes in the *S.cer x S.kud* hybrids mirror the *S.cer* subgenome counterparts (**Figure 1, Figure S1C**).

The phylogenetic structure of the hybrid subgenomes offers insight into the origins of these interspecific hybrids. First, the *S.cer* subgenomes of the lager beer lineage stems from Beer 1

whereas the traditional Belgian beer lineage forms a sister clade to Beer 2, indicating that both major domesticated beer lineages were involved in the emergence of interspecific beer yeast hybrids. Secondly, the monophyletic clustering of each beer hybrid type suggests that the present-day strains are the result of only one hybridization event per hybrid clade, or few events involving very similar strains[53,56–59]. This implies that the present-day diversity is largely due to the spread and diversification of existing hybrids rather than multiple, independent emergence and selection events. Third, the *S.cer* progenitors seem to come from industrial niches closely associated with the ones from which the hybrids were isolated, suggesting that these successful industrial yeast hybrids formed close to the fermentation environments in which they are now found.

The *S.eub* subgenomes of the *S.eub x S.uva* hybrids form a monophyletic sister clade to those of the *S.cer x S.eub* lager strains, indicating that the *S.eub* parents were closely related. In contrast to the monophyly of the *S.eub* subgenomes, the *S.uva* subgenomes are genetically clearly separated, indicating that at least three hybridisation events gave rise to the *S.eub x S.uva* hybrids (**Figure S1D** and **Figure S2**). The *S.eub* subgenomes clearly separate according to geographical isolation (Germany vs. Belgium), suggesting that a single or few closely related *S.eub* strains formed multiple hybrids, which then evolved and diverged independently. The genetic diversity across the *S.uva* subgenomes is significantly higher than across the corresponding *S.eub* subgenomes (average nucleotide diversity π=2.11E-03 and π=4.66E-04 respectively, one-sided Mann–Whitney U test, $P < 2.2 \times 10^{-16}$) and there is no clear niche substructure, suggesting that *S.uva* strains move more freely across environments. Indeed, in contrast to *S.eub*, *S.uva* has been isolated all over the world from a wide array of niches, natural and man-made[60]. Moreover, the species is a known contaminant in brewing environments[37,61,62]. Coupled with the fact that all of the *S.eub x S.uva* hybrids were isolated from spontaneous fermentations, old bottles or brewing equipment, it is likely that the hybridisations between *S.eub* strains and *S.uva* contaminants occurred within the brewing environment. All *S.uva* hybrid subgenomes are also closely related to European strains (**Table S1**), and thus the hybrids likely originated in Europe. However, as the population structure of *S.eub* and *S.uva* is not described as elaborately as that of *S.cer*, an Asian origin (as suggested in[23]) cannot be formally ruled out.

Interspecific hybrid genomes are hallmarked by significant ploidy variation and chimeric chromosomes

Newly formed hybrids experience extensive genome reorganization resulting in aneuplodies and chimeric chromosomes[63,64]. We found significant variation within and between hybrid types in overall ploidy, copy number of large chromosomal fragments and full chromosomes, as well as in the degree of parental species contribution to the hybrid genome **(Figure 2)**.

The *S.cer x S.kud* hybrids are overall triploid, with diploid *S.cer* and haploid *S.kud* subgenomes (**Figure 2A**). Notable exceptions are two Belgian spontaneous beer fermentation strains, ABI1606 and ABI1525, with a total ploidy of roughly 4n (triploid *S.cer*, haploid *S.kud*).

The *S.uva* x *S.eub* strains are divided into several subgroups with respect to subgenome content **(Figure 2B)**. The first subgroup, containing a subset of Belgian Lambic strains, exhibits a diploid *S.uva* subgenome and an extremely fragmented *S.eub* subgenome. The second subgroup, consisting of German brewing contaminant strains, presents a more uniform 1:1 *S.uva:S.eub* ratio, with multiple partial or complete chromosome deletions and a few duplicated chromosomal segments. The last subgroup contains two Lambic strains ABIC1571 and ABIC1602 that exhibit an intermediate genome composition (**Figure 1, Figure S2**).

In *S.cer x S.eub* strains, there is a clear distinction in genomic composition between the Saaz and Frohberg lineages (**Figure 2C and 2D**). Saaz strains are typically triploid (haploid *S.cer*, diploid *S.eub*)[21,53]. Frohberg strains are generally tetraploid to pentaploid with a basal 2n:2n ratio of the parental subgenomes. This high ploidy level is also in line with previously reported Frohberg genomes[20,33,53,56,59,64,65]. Both types demonstrate severe deletions and amplifications of large segments and even full chromosomes; Saaz strains mostly harbour losses in *S.cer* and amplifications of *S.eub* whereas Frohberg strains demonstrate losses in *S.eub* and amplifications of *S.cer*.

Most chromosomal regions exhibit integer ploidy changes but we did find a few instances of strains within each hybrid type with intermediate ploidy changes (e.g. regions of ~0.5 ploidy increments). Given that sequencing was performed on pure culture stocks that underwent a single cell bottleneck, we would have expected relatively isogenic populations. However, intermediate ploidies suggest that the hybrids genomes are unstable and that the populations

used for isolating genomic DNA carried unfixed genomic rearrangements. Interestingly, when further investigating the instability using PCRs targeting regions identified as unstable in four strains (ploidy between 0 and 1), we could only confirm instability in one strain, namely BE138 (**Figure S3**). In this strain, part of ChrIV of the *S.cer* subgenome was present in some clones but absent in others, proving that the hybrid genome is indeed unstable. We did not detect instability of the targeted regions in the other three strains using PCR, which may indicate that the patterns of (partial) chromosome losses are not necessarily the same in different strain subpopulations, such as the subpopulations used for PCR and the subpopulations used for genome sequencing.

Within the same hybrid type, we also observed striking differences in the copy number of full chromosomes and large chromosomal fragments. These copy number changes are often shared by some but not all strains originating from the same hybridisation event, and copy number differences occur even among closely related strains, indicating that post-hybridisation genome structural rearrangements are still ongoing (**Figure 2**). In many cases, a copy number change in one subgenome is compensated by an opposite copy number change on the homoeologous portion of the other subgenome, suggesting the large-scale occurrence of genomic rearrangements leading to chimeric chromosomes and often to loss of heterozygosity (LOH).

Mapping of these chimeric regions revealed more than 300 breakpoints amongst the hybrids (72 in *S.cer* x *S.kud*, 80 in *S.cer* x *S.eub*, and 150 in *S.uva* x *S.eub*, **Figure 3-ABC, Figure S4, Figure S5-ABCD**). The higher occurrence of breakpoints in *S.uva* x *S.eub* hybrids could be due to the lower nucleotide divergence between the two subgenomes (average whole-genome nucleotide identity 91.7%) compared to *S.cer* x *S.kud* (84.9%) and *S.cer* x *S.eub* (84.5%), which leads to a higher frequency of homologous recombination[14,58,59]. In fact, DNA sequence homology between subgenomes is significantly higher in breakpoint regions compared to non-breakpoint regions within stretches ranging from 50bp (microhomology) up to 1kb (**Figure S5-EFG**).

Breakpoint similarities and differences across the hybrid strains also offer an opportunity to further trace back their origin and evolutionary trajectory during diversification (**Figure 3-ABC**). Given the larger number of strains and higher coverage of the *S.cer* x *S.eub* hybrids, we focused our analysis on this group (**Figure 3C**). Out of the 80 identified breakpoints, 2 are found in all strains. An additional 4 breakpoints are found across all three subclades

(Frohberg, Saaz 1 and Saaz 2), 4 are found in both Saaz 1 and Saaz 2, 2 in Saaz 1 and Frohberg and 2 in Saaz 2 and Frohberg, but the majority (82.5%) are subclade-specific (33 Frohberg, 12 Saaz 1, 21 Saaz 2). The sharing of some breakpoints across the Frohberg and Saaz subclades supports a common origin of all lager yeasts. However, the exact trajectory and relationship of the subclades is difficult to disentangle, given that several breakpoints may have been present at one time but may have been obscured by chromosomal losses (e.g. the complete loss of *S.cer* chromosome XII in all Saaz strains). On the other hand, we cannot exclude the possibility that some of the breakpoint sites are more susceptible to rearrangements (fragile sites) and that shared breakpoints might have arisen independently in different hybrid lineages[53,56,59,65].

Hybrid beer yeasts exhibit unique phenotypic features that reflect niche adaptation

Large-scale changes in genome content and structure are intrinsically linked to phenotypic changes, which may confer fitness advantages[66–68]. To assess the extent of phenotypic changes in our hybrids, we extensively phenotyped the sequenced isolates and multiple pure species. Assays covered several industrially relevant traits including stress tolerances and metabolite production.

Based on overall phenotypic behaviour, strains cluster into three major groups, each correlated with a distinct genetic origin and industrial niche (**Figure 4A**). Group A contains all pure *S.cer* strains and *S.cer x S.kud* hybrids plus one *S.cer x S.eub* hybrid, which form several subgroups with distinct phenotypic profiles. One of the most distinguishing features between these subgroups is the division between beer-like traits and wine-like traits. Specifically, subgroups A3 and A4 (Trappist *S.cer x S.kud* hybrids and *S.cer* ale strains) exhibit weaker environmental stress resistance and sporulation efficiency than subgroups A1, A2, and A5 (strains used in wine, sake, spirits, cider, bioethanol, and Lambic brewing). This is indicative of a strong domestication signature (genome decay) in A3 and A4 which is common in ale strains[28,29]. Group B includes the non-*cerevisiae* pure species as well as all but two *S.uva x S.eub* hybrids. Except for the pure *S.kud* strains in subgroup B1, these strains demonstrate high cold, osmo-, and desiccation tolerance. Group C contains all but two lager hybrids and further subdivides in Frohberg (C1) and Saaz (C3). They can tolerate lower temperatures than strains from Group A. Tolerance to extreme cold (4°C) is limited compared to strains from Group B. Group C exhibits overall lower environmental stress tolerance than

yeasts from non-beer or spontaneous fermentation environments (primarily in Groups A and B).

Cold tolerance (a common trait in non-*cerevisiae* pure species) and the ability to efficiently use maltotriose (a beer-specific carbon source fermentable by many *S.cer* strains but typically not by non-*cerevisiae* species) are commonly combined in interspecific hybrids (**Figure 4B**). This combination has been proposed as the reason why these hybrids persisted and flourished in cold brewing environments[39,69,70] and could either be the result of post-hybridisation adaptation or a direct consequence of the hybridisation process itself. To directly evaluate these possibilities, we phenotyped several newly developed artificial *S.cer x S.eub* hybrids and compared them to their parental strains (**Figure 4C**). The artificial hybrids demonstrated improved cold tolerance compared to the *S.cer* parent (ANOVA $F=8.88$, $p < 0.001$) and improved maltotriose utilisation compared to the *S.eub* parent (ANOVA $F= 75.53$, $p < 0.001$), showing that hybridisation can generate immediate fitness advantages in niches such as cold beer fermentation.

Although hybridisation can impart immediate fitness advantages, the instability of interspecific hybrids likely facilitates further adaptation to specific beer niches, best exemplified by the 'traditional Belgian beer' lineage. This lineage likely originated from a single hybridisation event (see **Figure 1**) and subsequently split into two distinct subgroups, 'Lambic' and 'Trappist' strains, which function in two profoundly different beer production processes. During Lambic beer production, the presence of acid-producing bacteria leads to considerable concentrations of acetic and lactic acid[71,72], whereas in Trappist beer production, these bacteria are much less prominent. Despite their shared origin, tolerance to organic acids is observed in the 'Lambic', but not in the 'Trappist' subgroup, suggesting adaptation after the hybridisation event (**Figure 4D**).

Structural rearrangements underlie phenotypic convergence to low off-flavour production in lager yeast

Beer flavour and aroma diversity can often be attributed to the yeast (and, in some cases, bacteria) used during the fermentation process[73]. The style of beer dictates which specific aroma compounds are desirable and has thus influenced selection over time. The presence of 4-vinyl guaiacol (4-VG) is tolerated in some specialty beers such as wheat and saison beers, but undesirable in most other beer styles, including lagers. This compound produces a spicy, clove-like aroma referred to as 'phenolic off-flavour' (POF). Production of 4-VG depends on

functional copies of two subtelomeric genes, phenylacrylic acid decarboxylase (*PAD1*) and ferulic acid decarboxylase (*FDC1*)[74].

Our phenotype analysis demonstrated that all non-*cerevisiae* pure strains (*S.eub, S.kud,* and *S.uva*) produce 4-VG (=POF^+) (**Figure 5A**). As a dominant trait, newly formed hybrids with a non-*cerevisiae* parent should be POF^+ regardless of the phenotype of the *S.cer* parental strain. However, in contrast to all *S.uva x S.eub* and *S.cer x S.kud* hybrids, all *S.cer x S.eub* hybrids were found to be POF^- (**Figure 5A**), suggesting that 4-VG production was lost after hybridisation in lager strains. To understand the genetic foundation of this loss, we further evaluated the subgenomes of the lager hybrids.

The *S.cer* subgenomes of these strains localise within the Beer 1 clade; the majority of strains from this clade acquired disruptive mutations in *PAD1* and/or *FDC1*[28,29,75], suggesting that the hybrids inherited an inactive POF pathway from their *S.cer* ancestor. Our analysis shows that Frohberg strains do indeed harbour the same disruptive mutation(s) as other Beer 1 *S.cer* strains (**Figure 5B**). Saaz strains, however, harbour complete deletions of the *S.cer PAD1-FDC1* gene cluster (**Figure 5B**). Nevertheless, the hybrids must also have inherited an inactive POF pathway from the *S.eub* parent to render them POF^-. Three distinct genomic changes resulted in the loss of the *PAD1-FDC1* gene cluster in the *S.eub* subgenomes of the Saaz 1, Saaz 2 and Frohberg lineages (**Figure 5C**), each involving a different chimeric breakpoint between the terminal regions of *S.cer* ChrXIII and *S.eub* Chr13. Thus, each lineage experienced an independent LOH event, further supporting a post-hybridisation loss of 4-VG production, most likely after the divergence of Frohberg and Saaz. Interestingly, this was the only region in the *S.eub* subgenome that was lost three times independently in lager strains, signifying the strong selection against the POF phenotype.

Historical context of the origin and evolution of lager yeasts

Our data offer a unique opportunity to map evolutionary events in a historical perspective and gain insight into the potential driving forces behind yeast hybrid domestication. Absolute dating of evolutionary events in yeasts is complicated due to a lack of solid calibration points (for example, fossils) and the difficulty of tracing historical movements of industrial yeasts. Here, we used the split between UK and US *S.cer* strains in the Beer 1 clade, which is thought to have occurred between 1607 AD and 1637 AD when British settlers imported to US beer yeasts[28], as a calibration point to date divergences within and between the ale and lager subclades of Beer 1 (see **Methods**).

This calibration yields a timeline for beer yeast evolution that correlates with known historical events (**Box Figure 6** and **Supplementary Note 1**). We estimate that the most recent common ancestor of present-day ale and lager beer strains dates to the mid-16th century (**Figure 6**). The lager yeast lineage splits from the Belgium/Germany clade within the same time frame, shortly after the enactment of the Bavarian Beer Purity Law ('Bayerisches Reinheitsgebot') in 1516 AD. This law historically had a large, reductive impact on the variety of German beer styles as many local beers and brewing traditions disappeared in order to conform. The primary objective was to reduce bacterial contaminations by establishing rules regarding ingredients and restricting brewing to the colder winter months. Cold temperature brewing likely inadvertently selected for cold-tolerant strains – a trait characteristic of *S.eub*. This German origin of lager yeasts ties up with historical brewing records, which allowed us to trace back the origins of Saaz and Frohberg yeast to Bavaria, a region in the Southeast of Germany. A more elaborate note on lager beer yeast history can be found in **Supplementary Note 1**.

A notable difference between the lager and ale clades is the sudden, dramatic reduction of evolutionary rates of the lager lineages approximately 150 years after the split between the Saaz and the Frohberg lineage. The rate decrease correlates to the onset of diversification of both lineages, suggesting that the limited genetic diversity observed amongst today's commercial lager strains is a consequence of a genetic bottleneck and subsequent slow local divergence at the turn of the 20th century. This period coincides with three important events that revolutionised (lager) beer production. Firstly, Louis Pasteur described the importance of yeast in alcoholic fermentations in 1857 AD[76]. Once brewers became aware of the true significance of their sediment and the possible economic implications, greater care was taken to maintain their brewing cultures. Secondly, mechanical refrigeration was introduced into breweries from 1873 AD onwards[55,77]. This technological advance allowed for year-round lager fermentations as well as cooled storage of successful yeast cultures. Many brewers adopted the practice of keeping a separate, refrigerated population of their yeast, from which they regularly re-grew a larger population to inoculate a new brew. Additionally, mechanical cooling also allowed for brewing at colder temperatures, fostering selection for cold-tolerant strains. Colder fermentation temperatures and cold storage of yeast cultures partly explains the observed reduced evolutionary rate. Third, the isolation of the first pure yeast culture in 1883 AD (which was later shown to be a *S.cer x S.eub* hybrid) by E. C. Hansen at the Carlsberg brewery (Denmark) gave rise to a wave of brewers isolating and sharing their

strains[77]. The first isolated culture ('Unterhefe nr. 1') was disseminated across many breweries in central Europe, and rapidly implemented in their production process[78]. Together, these three events led to standardisation of industrial lager production and likely only a few closely related lager hybrid yeasts were disseminated across different breweries, where they were mostly preserved and used in cooled environments, slowing down their evolutionary divergence and further enhancing the selection for cold-tolerance.

Discussion

Industrial, man-made environments challenge microbes with unique environmental conditions and therefore offer insight into the processes that allow colonization of novel ecological niches. Our results show that interspecific hybridisation is an important and common route towards diversification and adaptation to novel niches. Moreover, our genome analyses and phenotyping suggest that adaptation is likely fuelled both by the direct selective advantages of the new hybrids, as well as their genomic plasticity, allowing for swift adaptation to specific niches. Beer yeasts provide a perfect example of this phenomenon, as interspecific hybridisation yielded new variants that combine the fermentation capacity of *S.cer* with the cold tolerance of other species. The instability and plasticity of the hybrid genomes likely allowed further adaptation and yielded variants that lost undesirable properties, such as the production of 4-vinyl guaiacol.

Our discovery of a large proportion of interspecific hybrids in Belgian specialty beers suggests that interspecific hybridisation played an important role in the history of industrial fermentations. Over the last century, the adoption of new brewing technologies and highly controlled single-strain fermentation processes likely contributed to the decline of the natural beer yeast diversity, including interspecific hybrids that may have been present as low-number contaminants of most ale fermentations. The low diversity of present-day lager yeasts exemplifies this trend. Although the first *S.cer x S.eub* hybrids originated around the 16th century AD, today's lager yeasts can be traced back to only a few lineages that were pure-cultured, cold-stored, and dispersed across multiple breweries in the late 19th century. Belgium represents a notable exception to this biodiversity decline. Traditional Belgian beer styles harbour a remarkably diverse array of yeasts, likely as a result of the continued use of old beer brewing practices. For example, the production of Belgian Lambic beers has remained unchanged for centuries. The use of medieval brewing technologies such as open-air inoculation and long-term fermentation and storage in barrels stored in the brewery's

cellar promoted the survival of unique yeast hybrids presented in this study, and could therefore provide a source of new biodiversity for industrial applications.

The study of the genetic and phenotypic makeup of naturally occurring yeast hybrids may further aid in the development of innovative new hybrids for industry. Recent studies have shown that recreation of interspecific hybridisation events that occurred in nature or development of new combinations of species results in hybrids with a phenotypic landscape beyond that of the strains usually employed in industrial fermentations[63,79,80]. Some of these hybrids produce new aromas that cater to changing consumer demands and demonstrate superior performance in challenging production environments. Detailed insight into how these hybrids evolved in different niches and which phenotypic features are retained from the parental species can aid in selecting suitable strains for development of industrially relevant hybrids.

Acknowledgements

We thank all K.J.V. and S.M. laboratory members for their help and suggestions. We thank AB InBev, White Labs Inc. and R. Steven Wagner (Central Washington University) for isolating and sharing various hybrid yeasts from the beer environment. We thank Martin Zarnkow (Research Center Weihenstephan for Brewing and Food Quality) for sharing historical facts on lager brewing. B.G. acknowledges funding from the VIB International PhD Program in Life Sciences. J.S. acknowledges funding from Research Foundation-Flanders (FWO) grant 12W3918N. Research in the lab of K.J.V. is supported by KU Leuven Program Financing, European Research Council (ERC) Consolidator Grant CoG682009, Human Frontier Science (HFSP) program grant RGP0050/2013, VIB, European Molecular Biology Organization (EMBO) Young Investigator program, FWO and VLAIO (Vlaams Agentschap Innoveren en Ondernemen). Research in the lab of S.M. is supported by VIB, Ghent University and FWO grant G018915N.

Contributions

Conceptualization: B.G., J.S., S.Maere, K.J.V; Data collection: B.G., J.S.; Formal analysis: B.G., J.S., G.B.; Investigation: B.G., J.S., S.M., J.L.G., R.W., F.A.T., F.B., V.S., B.H.-M, G.B.; Data visualisation: B.G., J.S.; Resources: M.H., F.M., P.M., B.S., L.D., T.P., C.W, K.J.V; Writing: B.G., J.S., M.C.D., S.Maere, K.J.V; Supervision: S.Maere, K.J.V.

Conflict of interests

Funding: AB InBev partially funded the experiments performed in this study. AB InBev had no role in conceptualization, design, data collection, analysis, decision to publish or preparation of the manuscript. **Employment:** Several authors are employed by AB InBev. BS, PM, LD.

Data availability

All raw sequencing reads generated in this study have been deposited to the NCBI Short Read Archive (BioProject accession:PRJNA516557).

Figure legends

Figure 1. Tanglegram depicting the relationships between *Saccharomyces* pure species and interspecific hybrids. Each line of the tanglegram represents one hybrid sequenced in this study, coloured by hybrid-type. Species subgenomes within the same hybrid strain are connected between the phylogenetic trees of (a) *Saccharomyces cerevisiae*, (b) *Saccharomyces eubayanus*, (c) *Saccharomyces uvarum*, and (d) *Saccharomyces kudriavzevii*. Lineages that do not contain hybrid strains are collapsed (diamonds). Expanded trees including strain origins are reported in **Figure S1**.

Figure 2. Genome structure of *Saccharomyces* interspecific hybrids. Ploidy profiles of (a) *S.cer x S.kud*, (b) *S.uva x S.eub*, (c) *S.cer x S.eub* (Saaz), and (d) *S.cer x S.eub* (Frohberg) hybrids. Top panels: strains are sorted based on the phylogeny of the *S. cerevisiae* or *S. uvarum* parental species (see **Figure S1**). Chromosomes are coloured based on calculated ploidy. Density plots to the left of each tree represent the per-species ploidy distribution aggregated across all the strains of each hybrid type. Bottom panels: detailed representation of the genomic contribution of the two parental species in one selected hybrid.

Figure 3. Distribution of chimeric breakpoints across interspecific hybrids. Presence (red) and absence (white) of specific chimeric breakpoints are shown for (a) *S.cer x S.kud*, (b) *S.uva x S.eub,* and (c) *S.cer x S.eub* hybrids. Strains (rows) are sorted phylogenetically according to the *S. cerevisiae*, *S. uvarum*, and *S. cerevisiae* subgenomes, respectively. Breakpoints (columns) are hierarchically clustered based on their presence or absence across strains. Strains for which low sequencing coverage level negatively affected the detection of breakpoints are indicated with an asterisk. Strain origins are colour coded per panel.

Figure 4. Trait variation and niche adaptation of interspecific hybrids. (a) Hierarchically clustered heat map of phenotypic diversity within interspecific hybrids and pure species. Phenotypic values are calculated as normalised z-scores. Missing values are shown in grey. Phenotypes (rows) are sorted based on five major categories (labelled at the left). (Sub)genome compositions are indicated at the branch tips and coloured by species. (b) Correlation of cold tolerance vs. maltotriose utilization of hybrids (circles) and pure species (triangles). (c) Performance of artificial interspecific hybrids and parental species in beer wort fermentations at cold temperatures. Three independent crosses of the same parents were performed (H1, H2, H3). Bars indicate mean ± sd of four biological replicates. Statistical significance determined by ANOVA between hybrids and the inferior pure species: *(P-value

≤ 0.05), **(P-value ≤ 0.01), ***(P-value ≤ 0.001). (d) Acid tolerance of *S.cer x S.kud* hybrids (circles) and pure species (triangles). Hybrids are coloured according to isolation origin.

Figure 5. The genetic basis of loss of 4-VG production in *S.cer* x *S.eub* hybrids. (a) Percentage of strains from each species or hybrid type that demonstrate production of 4-vinyl guaiacol (4-VG$^+$). (b) The genetic basis for loss of 4-VG production is depicted for the *S. cerevisiae* subgenomes (left) and *S. eubayanus* subgenomes (right) arranged by hybrid subgroup (shown sorted phylogenetically on the left). The subtelomeric position of the *PAD1-FDC1* gene cluster (red triangle) is shown on *S. cerevisiae* chromosome IV and *S. eubayanus* chromosome 13. Black arrows indicate the location of chimeric breakpoints between the homoeologous chromosomes. Chromosomes are coloured according to the ploidy of one representative strain (Saaz group 1: BE137, Saaz group 2: CBS1538, Frohberg: BE104). The loss-of-function mutation in the Frohberg *S.cer* subgenome is highlighted in the tan triangle.

Figure 6. Time-calibrated phylogeny of *S. cerevisiae* Beer 1 clade. The five subclades of Beer 1 are indicated to the right of the tree. Shaded vertical boxes highlight major splits within the *S.cer* Beer1 clade: split of present day ale and lager yeasts (yellow), split of European and US ale yeasts (green); split of Saaz and Frohberg lager yeasts (light blue). Branches are coloured according to their branch-specific average number of substitutions/site/year. Node bars represent 95% Highest Posterior Density (HPD) intervals. Dates on the top indicate relevant events in lager beer history listed in the Historical Data box on the right (see **Supplementary Note 1**).

Methods

gDNA Extraction

For strains BE114, BE116-BE130, BE132-BE136 and SP012 genomic DNA was extracted with the MasterPureTM Yeast DNA Purification Kit (Epicenter, USA). For the other strains, genomic DNA was prepared using the GENTRA PUREGENE Yeast KIT (Qiagen, Germany) with some modifications to the recommended protocol. The main modification involves a 2-hour treatment of the overnight cell culture with zymolyase to efficiently digest yeast cell wall. Final DNA concentrations were measured using Qubit (Thermo Fisher Scientific, USA), 260/230 and 260/280 ratios with Nanodrop (Thermo Fisher Scientific, USA).

Library preparation and Whole-genome sequencing

For strains BE114, BE116-BE130, BE132-BE136 and SP012 paired-end sequencing libraries were prepared using the Nextera XT DNA Library Preparation Kit. Sequencing was performed on a HiSeq 2500 system at Illumina (San Diego, USA). For the other strains paired-end sequencing libraries (MiSeq reagent kit v3, 600 cycles) with a mean insert size of ~300bp were prepared and run according to the manufacturer's instructions on an Illumina MiSeq at the Nucleomics Core facility in Leuven (http://www.nucleomics.be/).

Reference-based alignments and variant calling

Reads were pre-processed by filtering low quality and ambiguous reads, adapters and PhiX contaminations, using Trimmomatic (v0.30). Clean reads were mapped to 7 *Saccharomyces* species (*Saccharomyces* species complex): *Saccharomyces cerevisiae* reference genome S288c (R64-1-1, EF4-Ensemble Release 74), *Saccharomyces paradoxus* (strain YP138)[32], *Saccharomyces mikatae* (strain IFO 1815)[81], *Saccharomyces kudriavzevii* (IFO 1802)[81], *Saccharomyces arboricola* (strain H-6, NCBI:txid1160507)[82], *Saccharomyces eubayanus* (strain FM1318)[56], *Saccharomyces uvarum* (CBS 7001)[81] with the Burrows-Wheeler Aligner (BWA v0.7.17 aln)[83] using default parameters except for –q 10. Non-primary alignments were filtered out and duplicate reads were marked using Picard Tools (v1.56) (http://picard.sourceforge.net). Coverage was estimated based on read depth in non-overlapping 1kb windows (reported mean coverage per window) using BEDtools (v2.27.0)[84].

De novo assembly

For each library, low quality and ambiguous reads were trimmed using Trimmomatic (v0.30). Reads were error-corrected and subsequently used for the de novo assembly with SPAdes (v3.10.1)[85]. Next, the redundans pipeline (v1.2)[86] was used to detect and remove redundant contigs and perform scaffolding and gap closing based on paired reads information. In order to determine the coordinates of contigs from each newly assembled strain relative to the reference strain and to obtain pseudo-chromosomes, whole genome alignments were performed against the species identified in 1.3 using Ragout (v1.2)[87].

Phylogenomic analyses

In order to infer the origin of the hybrid *Saccharomyces* genomes and their genetic relationship across species and strains within the *Saccharomyces* species complex, genes that are orthologs and present in exactly one copy among strains and across species have been identified (single-copy orthologs). The starting set of genes included 4,722 1:1:1:1 orthologs among *S.cerevisiae* s288c, *S.paradoxus*, *S.kudriavzevii* IFO1802 and *S.uvarum* CBS7001 identified by Scannell and co-authors[81]. The starting set of genes was reduced to 4,125 genes after including *S.eubayanus* FM1318[56] and the outgroup species *Saccharomyces castellii*[88]. Next, the presence of these genes and their single-copy status was tested within a collection of 420 *Saccharomyces* isolates. The collection of strains investigated included 283 *S. cerevisiae isolates*, 3 *S. kudriavzevii* isolates, 43 *S. uvarum* isolates, 21 *S. eubayanus* isolates, 10 *S. cerevisiae x S. kudriavzevii* hybrids, 46 *S. pastorianus* hybrids and 13 *S. uvarum x S. eubayanus* hybrids (**Table S1**). From this step onward, the subgenomes of the hybrid isolates were considered as distinct species: for instance, for a *S. uvarum* x *S. eubayanus* hybrid, two sets of orthologs were identified, one for each species respectively. A local BLAST database was set up for all the genomes based on their *de novo* assembly (collapsed representation of each species subgenome) and BLASTN searches were performed (1E-04 E-value cut-off, >98% similarity and >85% coverage – blast v2.5.0+)[89,90] using for each species the set of genes identified in the previous step. For hybrid genomes an additional BLASTN step was implemented to compare the set of genes identified for the distinct parental species and to exclude genes with high similarity between species that cannot be unequivocally assigned to one or the other species. Five sets of genes were obtained: 1) 1,389 genes across the *S. cerevisiae* genomes and subgenomes, dubbed the *"S.cer"* set; 2) 1,571 genes across the *S. eubayanus* genomes and subgenomes ("*S.eub*" set); 3) 1,364 genes across *S. uvarum* genomes

and subgenomes (*"S.uva"* set); 4) 1,750 genes across the *S. kudriavzevii* genomes and subgenomes (*"S.kud"* set). Considering the high level of species-specific subgenome loss and fragmentation observed in some *Saccharomyces* hybrids and in order to maximize the number of isolates included, 3% of missing data per gene was allowed. Extreme cases, with >50% missing genes per strain were excluded from the analysis (only ~200 *S. eubayanus* genes could be annotated for ABI1605 and it was therefore excluded from the *S. eubayanus* phylogeny). Multiple nucleotide sequence alignments (MSAs) for each gene in each set identified above were generated using MAFFT (v7.187)[91], with default settings and 1,000 refinement iterations. The MSAs were concatenated into supermatrices for each species using FASconCAT (v1.0)[92]. Quality checks and format conversions were performed using trimAl (v1.2)[93]. The final *S.cer* supermatrix included 337 taxa and 1,556,065 positions, 97.345 % nucleotides, 2.655% gaps and 0% ambiguities. The final *S.eub* supermatrix included 81 taxa and 2,323,546 positions, 88.141% nucleotides, 11.859% gaps and 0% ambiguities. The final *S.uva* supermatrix included 55 taxa and 1,453,393 positions, 91.772% nucleotides, 8.228% gaps and 0% ambiguities. The final *S.kud* supermatrix included 14 taxa and 2,635,158 positions, 90.741% nucleotides, 9.259% gaps and 0% ambiguities. Within each supermatrix, each gene was considered as a separate data partition. Twenty-five completely random starting trees for the *S.cer* supermatrix and 20 random starting trees for the *S.eub*, *S.uva* and *S.kud* supermatrices were obtained using RAxML (v8.2.8)[94]. Maximum-likelihood (ML) tree searches were performed on each fully random starting tree under the GTRGAMMA model (4 discrete rate categories) using ExaML (v3.0.17)[95] and the rapid hill climbing algorithm (-f d). During the ML search, the alpha parameter of the model of rate heterogeneity and the rates of the GTR model of nucleotide substitutions were optimized independently for each partition. The branch lengths were optimized jointly across all partitions. For each starting tree, the best tree was selected based on the highest log-likelihood score. Parameters and branch lengths were re-optimized on the best topologies with ExaML (-f E) using the median of the four rate categories for the discrete approximation of the GAMMA model of rate heterogeneity (-a). The tree with the best overall log-likelihood score of all tree inferences was considered the final ML tree. Non-parametric bootstrap analysis was performed on the concatenated matrices using RaxML. The a posteriori boot-stopping criterion[96] (MR bootstrapping convergence criterion) was applied to define the number of replicates. After every 50 replicates, the set of bootstrapped trees generated so far was repeatedly (1,000 permutations) split in two equal subsets, and the Weighted Robinson-Foulds (WRF) distance was calculated between the majority-rule consensus trees of both subsets (for each

permutation). Low WRF distances (< 3%) for >= 99% of permutations were used to indicate bootstrapping convergence. Convergence was reached after 200 replicates for the *"S.cer"* phylogeny: average WRF = 2.04%, percentage of permutations in which the WRF was <=3.00 = 99.8%; 200 replicates for the *"S.eub"* phylogeny: average WRF = 1.40%, percentage of permutations in which the WRF was <=3.00 = 99.7%; 600 replicates for the *"S.uva"* phylogeny: average WRF = 1.36%, percentage of permutations in which the WRF was <=3.00 = 99.2%; 50 replicates for the *"S.kud"* phylogeny: average WRF = 0.26%, percentage of permutations in which the WRF was <=3.00 = 100%. The final trees were visualized and rooted in R (v3.4.1)[97] with the ggtree package (v1.8.2)[98] using *S. paradoxus* as outgroup for the "*S.cer"* tree and *S. castellii* for the other trees.

Divergence time estimation

We used BEAST (v1.10)[99] to estimate divergence times in the "Beer 1" clade of the *S.cer* phylogenetic tree, using the topology of best scoring ML *S.cer* tree obtained for a supermatrix of 1,389 protein coding genes, as described in the previous section. Mosaic strains, harbouring mixed genetic backgrounds (e.g. Hefeweizen isolates), were excluded from the analysis (population structure analysed with fastStructure[100] v.1.0). To date the phylogeny we used a calibration prior on the split between US and Britain beer yeasts, using a normal prior with a 99% confidence interval falling between 1607 AD and 1637 AD, as based on historical events[28]. We assessed the performance of several molecular clock models using BEAST (v1.10)[99] in combination with BEAGLE 2.1.2[101]. Specifically, we analysed the data using a strict clock, an uncorrelated relaxed clock with an underlying lognormal distribution[102], a random local clock[103] and a fixed local clock with predefined clades (Lager-Frohberg, Lager-Saaz, Ale-Britain, Ale-Belgium/Germany, Ale-US)[104]. A pure-birth Yule speciation prior and a random starting tree were used for the Bayesian inference analyses through Markov chain Monte Carlo. Each analysis was run until ESS values of at least 100 could be obtained for all relevant parameters, as computed by Tracer 1.7[105]. Of these models, the random local clock provided a significantly better fit to the data than the three competing models, as estimated using generalized stepping-stone sampling[106]. The uncorrelated relaxed clock model yielded the lowest model fit to the data of the clock models tested, indicative of evolutionary rate shifts having occurred in the *S.cer* tree which the single rate distribution in the uncorrelated relaxed clock model is not able to account for. Additionally, the relaxed clock assigns a unique rate to every branch of a tree, but changes in the rate of evolution do not necessarily occur smoothly nor on every branch of a tree. Assuming strict clock rates within predefined

clades allows capturing (major) shifts in evolutionary rates between those clades, but does not allow for any rate variation within each clade. The random local clock on the other hand allows sampling the state space of all possible (strict) local clock models on all possible rooted trees and concluded that an estimated 50 rate changes occurred throughout the tree. Given that 50% of the prior probability within the random local clock model assumes no rate changes (and over 95% prior probability of less than three rate changes occurring), this shows that there is a strong signal in the data in favour of a large number of rate changes (more than those at the predefined clades in the fixed local clock model), which provides an additional argument for the random local clock significantly outperforming all other models. Given these findings, we present results for the random local clock model only, by summarizing its divergence time estimates in a maximum clade credibility tree using TreeAnnotator, which is part of the BEAST software package[99].

Estimation of ploidy and identification of chimeric regions across species

From the alignment of paired end reads on the multi-species reference genome, a new alignment file was generated for each strain by retaining paired reads with high mapping quality ($q > 20$), for which the two mates are mapping on chromosomes of the different subgenomes. E.g. for read pair A, one read is mapped on *S. cerevisiae* chr I and its mate on *S. eubayanus* chr 1. We dubbed these reads "discordant reads". The absolute number of discordant reads over the total number of reads was calculated in non-overlapping 1kb windows over the full multi-species reference genome excluding unplaced contigs and mitochondrial contigs. In order to identify windows in the genome supporting the presence of a chimeric event, we selected windows with at least 15 reads and a minimum of 3 discordant reads. This very conservative threshold allows the identification of potential chimeric events across areas of the genome with varying coverage levels within the same hybrid-type and across hybrids with different genome size and hence varying coverage levels. Windows with a minimum of 3 discordant reads were defined as "breakpoint" windows; breakpoint windows preceded and/or followed by another breakpoint window were defined "major breakpoint" windows, because of the presence of consecutive windows supporting the chimeric event. Since breakpoint windows often coincided with changes in ploidy, we simultaneously calculated the ploidy level and the occurrence of discordant reads in each 1kb window along the genome. First, the mean raw coverage per 1kb non-overlapping window as calculated with BEDtools (v2.27.0)[84] was smoothed using a running median function on windows of 11 consecutive 1kb windows using the CaTools package (v1.6, https://CRAN.R-

project.org/package=caTools) in R. Second, the density of the smoothed coverage was plotted per subgenome and for the full hybrid genome. In order to identify the mean value of smoothed coverage corresponding to the haploid ploidy level, a Gaussian mixture model was fitted to the density distribution and mean and standard deviation were calculated for each peak (peaks are a proxy for ploidy levels detected in the hybrid genome). Third, the smoothed raw coverage values were divided by the mean value of the haploid peak to obtain estimated ploidy values for each pre-computed window in the genome. Due to the presence of coverage depth noise and potential copy number heterogeneity in the population of sequenced cells, "in-between" (non-integer) ploidy levels were detected. In order to define integer ploidies and detect ploidy shifts, a second Gaussian mixture model was fitted on the distribution of estimated ploidy values for each strain. This allowed the identification of mean and standard deviation of each ploidy peak and the definition of "ploidy shift boundaries" based on the intersection points between the ploidy modes in the mixture model. The identification of breakpoint windows and their ploidy context was then followed by the identification of the exact mapping location of the reads within the breakpoint window across the two subgenomes for each hybrid strain based on the following steps: i) extract the reads from the selected breakpoint windows; ii) retain reads that are still paired after the identification and selection of breakpoint windows; iii) calculate for each pair their summed edit distance from the corresponding reference sequence, normalise it by the summed length of the two reads; iv) select reads with a percentage identity >=95% against the corresponding reference location; iv) intersect the position of the reads with annotated features from the corresponding reference sequence (**Table S2**). In order to calculate nucleotide percentage identity (%) between subgenomes for breakpoint windows and non-breakpoint windows, pairwise whole-genome alignments were obtained between *Saccharomyces* species corresponding to the species combination identified in our set of hybrids using Mugsy (v1.2.3)[107]. Next, the 1kb interval coordinates from the alignments against the reference sequences were mapped on the whole-genome alignments and nucleotide percentage identity was calculated as number of matches on the total amount of bases in the window excluding gaps. We repeated the calculations for 500,100 and 50bp windows in order to identify microhomology regions within the starting 1kb windows. Then, we compared the identity distribution across breakpoint windows and non-breakpoint windows. The nucleotide percentage identity distribution across breakpoint windows and non-breakpoint windows was compared using the Wilcoxon Signed Rank Test implemented in the MASS package (v7.3-47)[108] in R. Twenty-one strains were excluded from the ploidy estimation analysis due to a bias in the read depth

profile already observed and described in Gallone et al., 2016 and referred as "smiley pattern": for these samples coverage follows a convex trend with high depth at the terminal regions of the chromosomes that gradually decreases toward the centre (see Table S1 and [28]).

Investigation of population heterogeneity using PCR

To investigate whether the population heterogeneity observed in the ploidy profiles of some of the hybrids is caused by unfixed genomic rearrangements and losses in the populations concerned, we monitored the presence/absence of unstable regions, i.e. regions that exhibit a calculated ploidy level between 0 and 1, in populations of four strains (BE137, BE138, ABI160, BR005; 1 region per strain). For each unstable region, we also tested a stable region in its proximity for which no heterogeneity was observed. For each strain, we assessed 45 individual, randomly picked colonies. First, strains were streaked from the -80°C stock to standard agar plates (YPGlu 2% agar; Yeast Extract 1% w v^{-1}, Peptone 2% w v^{-1}, Glucose 2% w v^{-1}, agar 2% w v^{-1}) to single colonies. After a 2-day incubation at RT, the 45 random colonies were selected, and gDNA was extracted using a 10 min boil in NaOH. PCRs to assess the absence/presence of the target regions were performed (all primers are provided in **Supplementary Dataset 1**). The absence of a PCR product is indicative of loss of the region.

Flavour Production and Flocculation in Fermentation Conditions

To assess the metabolite production of the yeasts, lab-scale fermentation experiments were performed. These fermentations were performed in rich growth medium (YPGlu 10%; peptone 2% w v^{-1}, yeast extract 1% w v^{-1}, glucose 10% w v^{-1}) and beer wort (13°P, 8 EBC Brewferm, Belgium). Precultures were inoculated in test tubes containing 5mL of yeast extract (1% w v^{-1}), peptone (2% w v^{-1}) and glucose (4% w v^{-1}) medium (YPGlu 4%) and incubated overnight at 30°C (shaking). After 16 hours, the preculture was diluted 10-fold in 50 mL of YPGlu 4% medium and transferred to a 250 mL Erlenmeyer flasks. This second preculture was incubated for 16 hours at 30°C (shaking). Next, the preculture was used to incoulate the growth medium at an initial optical density at 600 nm (OD600) of 0.5 (roughly equivalent to 10^7 cells mL^{-1}). The fermentations were performed in 250 mL Schott bottles with a water lock placed on each bottle. They were incubated for 7 days at 20°C (YPGlu 10%) or 16°C (beer wort), statically. In addition, H_2S production during the fermentation was tracked using a lead acetate strip, which were scored from 0 (no colour reaction, white strip) to 3 (intense colour reaction, black strip) after the fermentation to quantify H_2S formation. To estimate fermentation progress, weight loss was measured daily. After 7 days, the

fermentations were stopped, filtered (using 0.15 mm paper filter) and samples for chromatographic, density, spectrophotometric and ethanol measurements were taken. Maltotriose utilization (%) in beer wort was calculated by comparing the total weight loss of the fermentation to the theoretical maximum (calculated based on total fermentable sugar concentrations). Additionally, after fermentation, the flocculation character of each strain was scored visually, using a score ranging from 1 (not flocculent) to 6 (extremely flocculent, big flocs).

Headspace gas chromatography coupled with flame ionization detection (HS-GC-FID) (Agilent Technologies, USA) was performed as described previously[28].

Acetic acid, sulfite, pH and glycerol production were analysed via the Gallery™ Plus Beermaster Discrete Analyzer (Thermo Fisher Scientific, USA), according to the manufacturer's recommendations.

Ethanol Accumulation Capacity

The maximal ethanol accumulation capacity of all strains was assessed as described previously[28].

Screening for Environmental and Nutrient Stress Tolerance

All strains were tested in several conditions using robot-assisted spotting assays. All strains were evaluated on YPGlu 2% agar for (i) temperature tolerance (4°C – 12°C - 16°C - 30°C - 37°C – 39°C), (ii) sugar- and/or osmotolerance using increasing concentrations of glucose (final osmolyte concentration of 44 - 46 - 48 w v^{-1}), (iii) acid tolerance using increasing concentrations of acetic acid (12,5 - 25 - 50 - 75 mM), (iv) ethanol tolerance using increasing concentrations of ethanol (5 - 7 - 9 - 10 - 11 % v v^{-1}), and (v) copper tolerance using 0.050-0,075-0,100 mM of copper. For each of these experiments, growth on YPGlu 2% agar on 20°C was used as a reference condition.

Spotting assays and image analyses were performed as described previously[28]. Heat maps were obtained using the R function heatmap.2 from gplots package (v3.0.1)[109]. Strains were hierarchically clustered based on phenotypic behaviour using ward.D2 method[110] on Euclidean distances.

Investigation of the Yeast's Sexual Life Cycle

Sporulation was induced on minimal sporulation medium [1% (w v^{-1}) KAc, 0.05% (w v^{-1}) amino acids, 2% (w v^{-1}) agar]. After pre-growth in 5mL YPGlu 2% (overnight at 20°C, shaking), strain were incubated at 23°C for 10 days. Dissection of 4 tetrads of each strain was carried out using a micromanipulator (Singer Instruments, UK), and mating-type determination of all segregants was performed by mating-type PCR.

Yeast survival in beer

To assess survival of all strains, the viability of cultures aging in Duvel Green [blond ale, 7% (v v^{-1}) ethanol] was tracked over a one month period. Yeast precultures were shaken for 48 hours at 30°C in test tubes containing 5 mL YPGlu 2%. This sample was then transferred to 15 mL falcons and centrifuged for 3 minutes at 3000 rpm. After discarding the supernatant, the samples were resuspended in sterile water to reach an initial cell count of $4*10^7$ cells mL^{-1}. 0,5 mL of this sample was used to inoculate a sterile GC vial containing 10 mL of filter-sterilized Duvel Green supplemented with 0.2% (w w^{-1}) glucose to reach a final concentration of $2*10^6$ cells mL^{-1}. The headspace of the samples was flushed with CO_2 before capping the vials. The vials were incubated statically at 30°C. After 30 days, samples were taken, and viability was assessed using a methylene blue [0.1% (w v^{-1}), Sigma-Aldrich] staining. Automated cell counting was performed with the TC20 automated cell counter (Biorad, USA).

Desiccation tolerance

Desiccation tolerance was measured by a modified version of the assay described in Calahan et al. 2011[111]. All strains were streaked on YPGlu 2% agar. Strains were pregrown in YPGlu 2%, supplemented with 0.5% v v^{-1} Tween80 and 20 µg mL^{-1} ergosterol in deep-well plates at 16°C for 3 days. Cells were harvested by centrifugation (3000 rpm, 3 minutes) and supernatant was removed. 10^8 cells were resuspended in 1 mL assay buffer (x8 dilution of phosphate buffered saline: 17.1 mM NaCl, 0.338 mM KCl,, 1.25 mM Na2HPO4, 0.220 mM KH2PO4, pH 7.4), after which viability was checked using plating. 100 µL of this cell culture was transferred to a 96-well microtiter plate. The lid of the 96-well microtiter plate was lifted 1.7 cm (using 4 small pieces of cardboard) during the experiment to allow sufficient air flow. Plates were incubated for 7 days at 8°C. Afterwards, cells were resuspended in 200 µL assay buffer and viability was assayed using plating. Next, the ratio between the viability after desiccation and the viability before desiccation was calculated. These data was subsequently

792 binned to score desiccation tolerance (viability <1% = '1', viability 1-20% = '2', viability 20-
793 50% = '3', viability >50% = '4'). All strains were tested in biological duplicates.

794 **Biomass production in cold wort**

795 To assess biomass production in beer wort at very low temperatures strains were pregrown in
796 100 µL of YPGlu+Mal 2% (Yeast Extract 1% w v^{-1}, Peptone 2% w v^{-1}, Glucose 1% w v^{-1},
797 maltose 1% w v^{-1}) at 16°C for 2 days. Next, 5 µL of these cultures were transferred to 95 µL
798 wort (8°P, 8 EBC Brewferm, Belgium) in a 96-well microtiter plate and incubated at 8°C for
799 7 days on a microtiter plate shaking platform (Heidolph Instruments, Germany) at 600 rpm.
800 Optical density of the strains was assessed after 0, 3 and 7 days. Experiments were performed
801 in duplicate.

802 **Development and phenotypic evaluation of artificial hybrids**

803 Hybridisation was induced by placing single spores from both parental strains together with a
804 micromanipulator (Singer instruments MSM, UK) on YPGlu 2% agar, followed by visual
805 inspection of zygote formation after 6-8 hours of incubation at room temperature. Candidate
806 interspecific hybrids were purified by streaking on wort agar medium [12% w v^{-1} malt extract
807 (8 EBC Brewferm, Belgium) and 1.5% w v^{-1} agar]. Hybrids were confirmed through a
808 species PCR[79,112,113]. PCR-confirmed interspecific hybrids were streaked another three
809 consecutive times on 12°P wort agar prior to long term storage at -80°C to ensure strain
810 purity. Three independent hybrids were developed from the same two parents (BE014 and
811 WL024). Ploidy investigation of the hybrids showed that all hybrids were triploid, which is
812 expected given the ploidy of the parental strains, which were diploid (WL024) and tetraploid
813 (BE014). This results in segregants that are haploid and diploid, respectively.

814 These hybrids were tested for biomass during wort fermentation at cold temperature as
815 described in the previous section. The maltotriose that was still present after 7 days of
816 fermentation was measured using the Dionex system (ICS 5000+, Thermo Fisher Scientific,
817 USA). Fermentation samples were diluted in dH_2O, filtered (0.2 µm), and 10 µL were
818 injected into the system. A Carbopac PA20 column, kept at 30°C, was used, with a column
819 flow of 0.3 µl min^{-1}. 250mM NaOH (Eluent 1) and 500 mM CH_3COONa + 100 mM NaOH
820 (Eluent 2) were used as eluents.

821

References

1. Mallet, J., Besansky, N. & Hahn, M. W. How reticulated are species? *BioEssays* **38**, 140–149 (2016).

2. Abbott, R. *et al.* Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).

3. Shapiro, B. J., Leducq, J.-B. & Mallet, J. What Is Speciation? *PLOS Genet.* **12**, e1005860 (2016).

4. Nolte, A. W. & Tautz, D. Understanding the onset of hybrid speciation. *Trends Genet.* **26**, 54–8 (2010).

5. Baack, E. J. & Rieseberg, L. H. A genomic view of introgression and hybrid speciation. *Curr. Opin. Genet. Dev.* **17**, 513–8 (2007).

6. Rieseberg, L. H., Archer, M. A. & Wayne, R. K. Transgressive segregation, adaptation and speciation. *Heredity (Edinb).* **83**, 363–372 (1999).

7. Lamichhaney, S. *et al.* Rapid hybrid speciation in Darwin's finches. *Science* **359**, 224–228 (2018).

8. Rieseberg, L. H. *et al.* Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica* **129**, 149–65 (2007).

9. Rieseberg, L. H. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *Am. J. Bot.* **78**, 1218–1237 (1991).

10. Schwarzbach, A. E. & Rieseberg, L. H. Likely multiple origins of a diploid hybrid sunflower species. *Mol. Ecol.* **11**, 1703–15 (2002).

11. Pryszcz, L. P. *et al.* The Genomic Aftermath of Hybridization in the Opportunistic Pathogen *Candida metapsilosis*. *PLOS Genet.* **11**, e1005626 (2015).

12. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).

13. Greig, D. Reproductive isolation in *Saccharomyces*. *Heredity (Edinb).* **102**, 39–44 (2009).

14. Liti, G., Barton, D. B. H. & Louis, E. J. Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. *Genetics* **174**, 839–50 (2006).

15. Hou, J., Friedrich, A., De Montigny, J. & Schacherer, J. Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in *Saccharomyces cerevisiae*. *Curr. Biol.* (2014). doi:10.1016/j.cub.2014.03.063

16. Morales, L. & Dujon, B. Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. *Microbiol. Mol. Biol. Rev.* **76**, 721–39 (2012).

17. Naumov, G. I. *et al.* Natural polyploidization of some cultured yeast *Saccharomyces* sensu stricto: Auto- and allotetraploidy. *Syst. Appl. Microbiol.* (2000).

18. Zhang, H., Skelton, A., Gardner, R. C. & Goddard, M. R. *Saccharomyces* paradoxus and *Saccharomyces cerevisiae* reside on oak trees in New Zealand: evidence for migration from Europe and interspecies hybrids. *FEMS Yeast Res.* **10**, 941–947 (2010).

19. Stefanini, I. *et al.* Social wasps are a *Saccharomyces* mating nest. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 2247–51 (2016).

20. Nakao, Y. *et al.* Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res.* **16**, 115–29 (2009).

21. Walther, A., Hesselbart, A. & Wendland, J. Genome sequence of *Saccharomyces carlsbergensis*, the world's first pure culture lager yeast. *G3 (Bethesda).* **4**, 783–93 (2014).

22. Dunn, B. & Sherlock, G. Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res.* **18**, 1610–23 (2008).

23. 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).

24. Cavalieri, D., McGovern, P. E., Hartl, D. L., Mortimer, R. & Polsinelli, M. Evidence for *S. cerevisiae* fermentation in ancient wine. *J. Mol. Evol.* **57 Suppl 1**, S226-32 (2003).

25. Michel, R. & McGovern, P. Chemical evidence for ancient beer. *Nature* 145–150

(1992).

26. McGovern, P. E. *et al.* Fermented beverages of pre- and proto-historic China. *Proc. Natl. Acad. Sci.* **101**, 17593–17598 (2004).

27. Borneman, A. R. *et al.* Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet.* **7**, e1001287 (2011).

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

29. Gonçalves, M. *et al.* Distinct Domestication Trajectories in Top-Fermenting Beer Yeasts and Wine Yeasts. *Curr. Biol.* **26**, 2750–2761 (2016).

30. 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).

31. Peter, J. *et al.* Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* **556**, 339–344 (2018).

32. Yue, J.-X. *et al.* Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* (2017).

33. Dunn, B., Richter, C., Kvitek, D. J., Pugh, T. & Sherlock, G. Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res.* **22**, 908–24 (2012).

34. Duan, S. F. *et al.* The origin and adaptive evolution of domesticated populations of yeast from Far East Asia. *Nat. Commun.* **9**, 2690 (2018).

35. Ludlow, C. L. *et al.* Independent Origins of Yeast Associated with Coffee and Cacao Fermentation. *Curr. Biol.* **26**, 965–971 (2016).

36. Barbosa, R. *et al.* Multiple Rounds of Artificial Selection Promote Microbe Secondary Domestication—The Case of Cachaça Yeasts. *Genome Biol. Evol.* **10**, 1939–1955 (2018).

905 37. Libkind, D. *et al.* Microbe domestication and the identification of the wild genetic
906 stock of lager-brewing yeast. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14539–14544 (2011).

907 38. Eizaguirre, J. I. *et al.* Phylogeography of the wild Lager-brewing ancestor (
908 *Saccharomyces eubayanus*) in Patagonia. *Environ. Microbiol.* (2018).
909 doi:10.1111/1462-2920.14375

910 39. Gibson, B. R., Storgårds, E., Krogerus, K. & Vidgren, V. Comparative physiology and
911 fermentation performance of Saaz and Frohberg lager yeast strains and the parental
912 species *Saccharomyces eubayanus*. *Yeast* **30**, 255–266 (2013).

913 40. González, S. S., Barrio, E. & Querol, A. Molecular characterization of new natural
914 hybrids of *Saccharomyces cerevisiae* and *S. kudriavzevii* in brewing. *Appl. Environ.*
915 *Microbiol.* **74**, 2314–20 (2008).

916 41. González, S. S., Barrio, E., Gafner, J. & Querol, A. Natural hybrids from
917 *Saccharomyces cerevisiae*, *Saccharomyces bayanus* and *Saccharomyces kudriavzevii*
918 in wine fermentations. *FEMS Yeast Res.* **6**, 1221–1234 (2006).

919 42. Masneuf, I., Hansen, J., Groth, C., Piskur, J. & Dubourdieu, D. New hybrids between
920 *Saccharomyces* sensu stricto yeast species found among wine and cider production
921 strains. *Appl. Environ. Microbiol.* **64**, 3887–92 (1998).

922 43. Peris, D., Lopes, C. a, Belloch, C., Querol, A. & Barrio, E. Comparative genomics
923 among *Saccharomyces cerevisiae* × *Saccharomyces kudriavzevii* natural hybrid strains
924 isolated from wine and beer reveals different origins. *BMC Genomics* **13**, 407 (2012).

925 44. Peris, D., Pérez-Torrado, R., Hittinger, C. T., Barrio, E. & Querol, A. On the origins
926 and industrial applications of *Saccharomyces cerevisiae* × *Saccharomyces kudriavzevii*
927 hybrids. *Yeast* **35**, 51–69 (2018).

928 45. Lopes, C. A., Barrio, E. & Querol, A. Natural hybrids of S. *cerevisiae*×S. *kudriavzevii*
929 share alleles with European wild populations of *Saccharomyces kudriavzevii*. *FEMS*
930 *Yeast Res.* **10**, 412–421 (2010).

931 46. Peris, D. *et al.* The molecular characterization of new types of *Saccharomyces*
932 *cerevisiae* × S. *kudriavzevii* hybrid yeasts unveils a high genetic diversity. *Yeast* **29**,
933 81–91 (2012).

934 47. Le Jeune, C. *et al.* Characterization of natural hybrids of *Saccharomyces cerevisiae*
935 and *Saccharomyces bayanus* var. *uvarum*. *FEMS Yeast Res.* **7**, 540–549 (2007).

936 48. Gallone, B. *et al.* Origins, evolution, domestication and diversity of *Saccharomyces*
937 beer yeasts. *Curr. Opin. Biotechnol.* **49**, 148–155 (2018).

938 49. Steensels, J., Gallone, B., Voordeckers, K. & Verstrepen, K. J. Domestication of
939 Industrial Microbes. *Curr. Biol.* **29**, R381–R393 (2019).

940 50. Wendland, J. Lager Yeast Comes of Age. *Eukaryot. Cell* **13**, 1256–1265 (2014).

941 51. G Gibson, B. & Liti, G. *Saccharomyces pastorianus*: Genomic insights inspiring
942 innovation for industry. *Yeast* **32**, 17–27 (2015).

943 52. Peris, D. *et al.* Complex Ancestries of Lager-Brewing Hybrids Were Shaped by
944 Standing Variation in the Wild Yeast *Saccharomyces eubayanus*. *PLOS Genet.* **12**,
945 e1006155 (2016).

946 53. Okuno, M. *et al.* Next-generation sequencing analysis of lager brewing yeast strains
947 reveals the evolutionary history of interspecies hybridization. *DNA Res.* **23**, dsv037
948 (2016).

949 54. Dai, F. *et al.* Tibet is one of the centers of domestication of cultivated barley. *Proc.*
950 *Natl. Acad. Sci. U. S. A.* **109**, 16969–73 (2012).

951 55. Meußdoerffer, F. & Zarnkow, M. Das Bier : Eine Geschichte von Hopfen und Malz.
952 (2014).

953 56. Baker, E. *et al.* The Genome Sequence of *Saccharomyces eubayanus* and the
954 Domestication of Lager-Brewing Yeasts. *Mol. Biol. Evol.* **32**, 2818–2831 (2015).

955 57. Monerawela, C. & Bond, U. Recombination sites on hybrid chromosomes in
956 *Saccharomyces pastorianus* share common sequence motifs and define a complex
957 evolutionary relationship between group I and II lager yeasts. *FEMS Yeast Res.* **17**,
958 (2017).

959 58. Belloch, C. *et al.* Chimeric genomes of natural hybrids of *Saccharomyces cerevisiae*
960 and *Saccharomyces kudriavzevii*. *Appl. Environ. Microbiol.* **75**, 2534–2544 (2009).

961 59. 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**, 1–71 (2014).

60. Almeida, P. *et al.* A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum*. *Nat. Commun.* **5**, 4044 (2014).

61. Peŕez-Través, L., Lopes, C. A., Querol, A. & Barrio, E. On the complexity of the *Saccharomyces bayanus* taxon: Hybridization and potential hybrid speciation. *PLoS One* **9**, e93729 (2014).

62. Back, W. *Colour atlas and handbook of beverage biology. Colour atlas and handbook of beverage biology.* (Fachverlag Hans Carl, 2005).

63. Krogerus, K., Magalhães, F., Vidgren, V. & Gibson, B. Novel brewing yeast hybrids: creation and application. *Applied Microbiology and Biotechnology* **101**, 65–78 (2017).

64. van den Broek, M. *et al.* Chromosomal Copy Number Variation in *Saccharomyces pastorianus* Is Evidence for Extensive Genome Dynamics in Industrial Lager Brewing Strains. *Appl. Environ. Microbiol.* **81**, 6253–6267 (2015).

65. Monerawela, C. & Bond, U. The hybrid genomes of *Saccharomyces pastorianus* : A current perspective. *Yeast* (2017). doi:10.1002/yea.3250

66. Fisher, K. J., Buskirk, S. W., Vignogna, R. C., Marad, D. A. & Lang, G. I. Adaptive genome duplication affects patterns of molecular evolution in *Saccharomyces cerevisiae*. *PLOS Genet.* **14**, e1007396 (2018).

67. Selmecki, A. M. *et al.* Polyploidy can drive rapid adaptation in yeast. *Nature* **519**, 349–352 (2015).

68. Todd, R. T., Forche, A. & Selmecki, A. Ploidy Variation in Fungi: Polyploidy, Aneuploidy, and Genome Evolution. in *The Fungal Kingdom* **5**, 599–618 (asm Pub2Web, 2017).

69. Baker, E. C. P. *et al.* Mitochondrial DNA and temperature tolerance in lager yeasts. *Sci. Adv.* **5**, eaav1869 (2019).

70. Li, X. C., Peris, D., Hittinger, C. T., Sia, E. A. & Fay, J. C. Mitochondria-encoded genes contribute to evolution of heat and cold tolerance in yeast. *Sci. Adv.* **5**, eaav1848

(2019).

71. De Roos, J. & De Vuyst, L. Microbial acidification, alcoholization, and aroma production during spontaneous lambic beer production. *J. Sci. Food Agric.* (2018).

72. Steensels, J. & Verstrepen, K. J. Taming Wild Yeast: Potential of Conventional and Nonconventional Yeasts in Industrial Fermentations. *Annu. Rev. Microbiol.* **68**, 61–80 (2014).

73. Dzialo, M. C., Park, R., Steensels, J., Lievens, B. & Verstrepen, K. J. Physiology, ecology and industrial applications of aroma formation in yeast. *FEMS Microbiol. Rev.* **41**, S95–S128 (2017).

74. Mukai, N., Masaki, K., Fujii, T., Kawamukai, M. & Iefuji, H. *PAD1* and *FDC1* are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **109**, 564–9 (2010).

75. Mukai, N., Masaki, K., Fujii, T. & Iefuji, H. Single nucleotide polymorphisms of *PAD1* and *FDC1* show a positive relationship with ferulic acid decarboxylation ability among industrial yeasts used in alcoholic beverage production. *J. Biosci. Bioeng.* **118**, 50–5 (2014).

76. Pasteur, L. Etudes sur la bière ses maladies, causes qui les provoquent, procédé pour la rendre inaltérable, avec une théorie nouvelle de la fermentation. 387 (1876).

77. Teich, M. *Bier, Wissenschaft und Wirtschaft in Deutschland 1800-1914: ein Beitrag zur deutschen Industrialisierungsgeschichte*. (Böhlau, 2001).

78. Hansen, E. Recherches sur la physiologie et la morphologie des ferments alcooliques V. Methodes pour obtenir des cultures pures de *Saccharomyces* et de mikroorganismes analogues. *C. R. Trav. Lab. Carlsb.* **2**, 92–105 (1883).

79. Mertens, S. *et al.* A Large Set of Newly Created Interspecific *Saccharomyces* Hybrids Increases Aromatic Diversity in Lager Beers. *Appl. Environ. Microbiol.* **81**, 8202–8214 (2015).

80. Peris, D. *et al.* Hybridization and adaptive evolution of diverse *Saccharomyces* species for cellulosic biofuel production. *Biotechnol. Biofuels* **10**, 78 (2017).

81. Scannell, D. R. *et al.* The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the *Saccharomyces* sensu stricto Genus. *G3 (Bethesda).* **1**, 11–25 (2011).

82. Liti, G. *et al.* High quality de novo sequencing and assembly of the *Saccharomyces arboricolus* genome. *BMC Genomics* **14**, 69 (2013).

83. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–60 (2009).

84. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

85. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–77 (2012).

86. Pryszcz, L. P. & Gabaldón, T. Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Res.* **44**, e113–e113 (2016).

87. Kolmogorov, M., Raney, B., Paten, B. & Pham, S. Ragout-a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* **30**, i302–i309 (2014).

88. Wapinski, I., Pfeffer, A., Friedman, N. & Regev, A. Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**, 54–61 (2007).

89. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

90. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

91. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–80 (2013).

92. Kück, P. & Meusemann, K. FASconCAT: Convenient handling of data matrices. *Mol. Phylogenet. Evol.* **56**, 1115–1118 (2010).

93. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–3 (2009).

94. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–3 (2014).

95. Kozlov, A. M., Aberer, A. J. & Stamatakis, A. ExaML version 3: a tool for phylogenomic analyses on supercomputers. *Bioinformatics* **31**, 2577–9 (2015).

96. Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E. & Stamatakis, A. How Many Bootstrap Replicates Are Necessary? *J. Comput. Biol.* **17**, 337–354 (2010).

97. R Development Core Team, R. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing* **1**, 409 (2011).

98. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

99. Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, (2018).

100. Raj, A., Stephens, M. & Pritchard, J. K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**, 573–89 (2014).

101. Ayres, D. L. *et al.* BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. *Syst. Biol.* **61**, 170–173 (2012).

102. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

103. Drummond, A. J. & Suchard, M. A. Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* **8**, 114 (2010).

104. Worobey, M., Han, G.-Z. & Rambaut, A. A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* **508**, 254–7 (2014).

105. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).

106. Baele, G., Lemey, P. & Suchard, M. A. Genealogical Working Distributions for Bayesian Model Testing with Phylogenetic Uncertainty. *Syst. Biol.* **65**, 250–264 (2016).

107. Angiuoli, S. V. & Salzberg, S. L. Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics* **27**, 334–342 (2011).

108. Venables, W. N., Ripley B.D. Modern Applied Statistics With S. *Technometrics* **45,** 111–111 (2003).

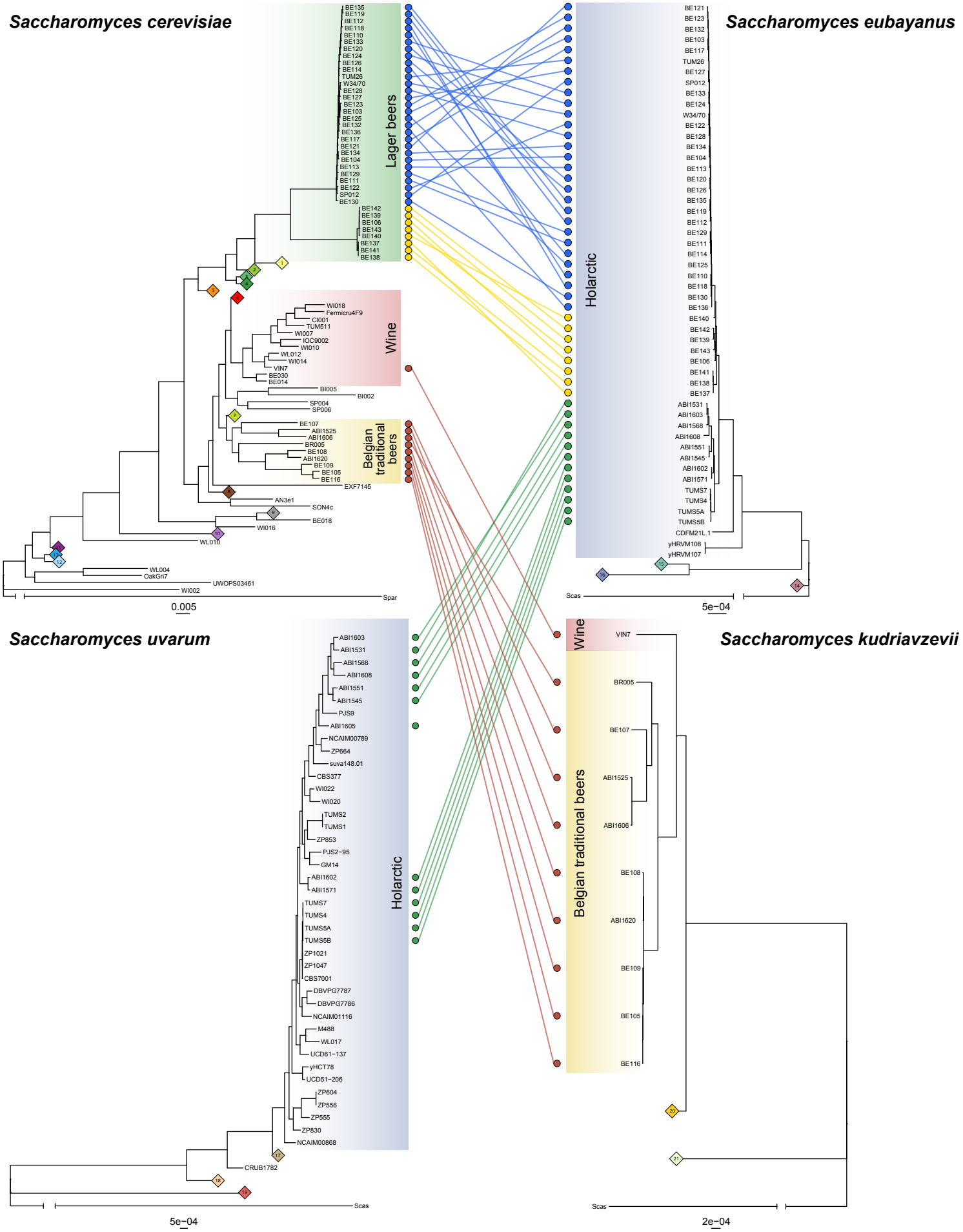
109. Warnes, G. R. *et al.* gplots: Various R Programming Tools for Plotting Data. R package version 3.0.0. http://CRAN.R-project.org/package=gplots. (2016).

110. Murtagh, F. & Legendre, P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *J. Classif.* **31**, 274–295 (2014).

111. Calahan, D., Dunham, M., Desevo, C. & Koshland, D. E. Genetic analysis of desiccation tolerance in *Saccharomyces cerevisiae*. *Genetics* **189**, 507–519 (2011).

112. Muir, A., Harrison, E. & Wheals, A. A multiplex set of species-specific primers for rapid identification of members of the genus *Saccharomyces*. *FEMS Yeast Res.* **11**, 552–563 (2011).

113. Pengelly, R. J. & Wheals, A. E. Rapid identification of *Saccharomyces eubayanus* and its hybrids. *FEMS Yeast Res.* **13**, 156–161 (2013).

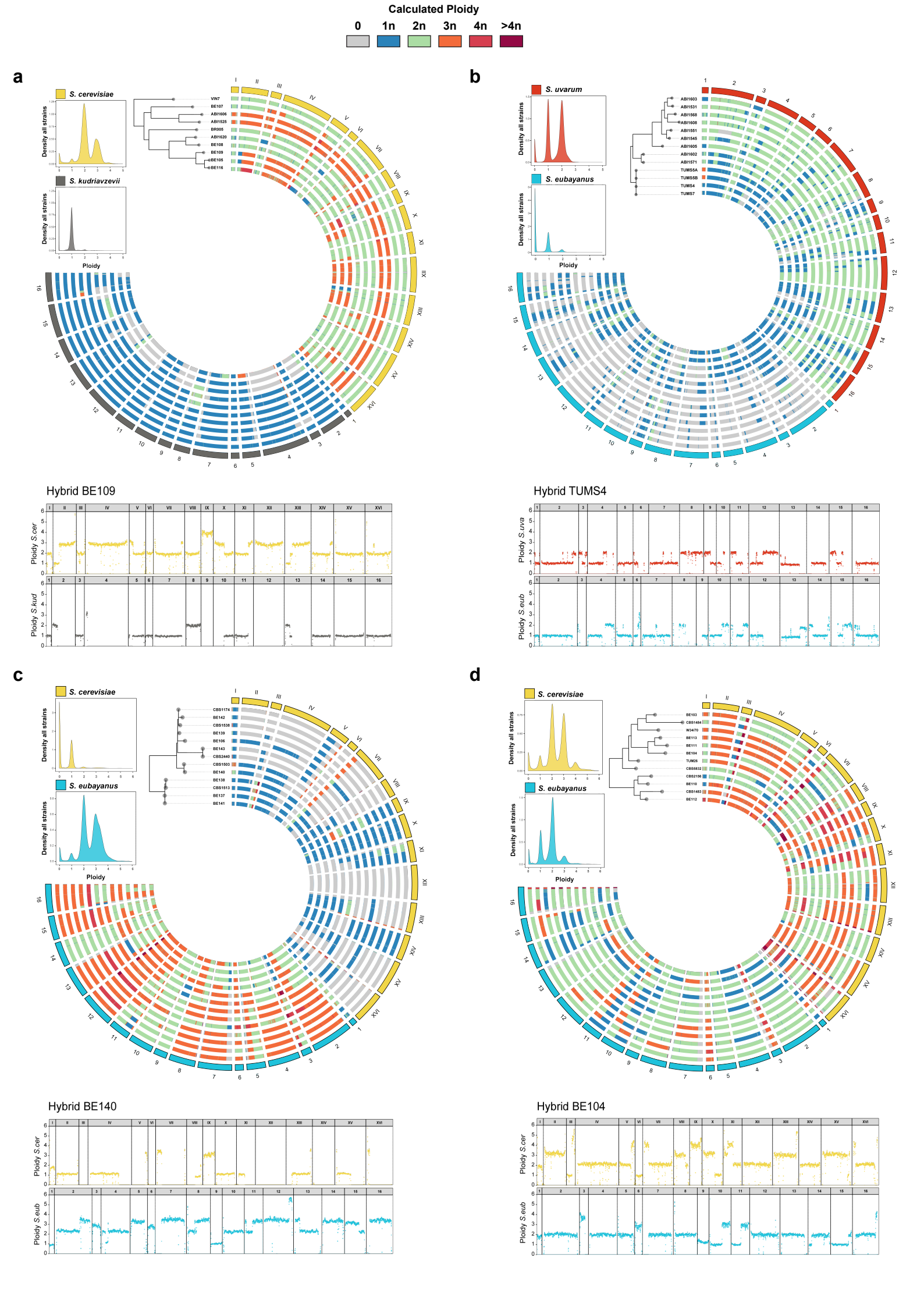


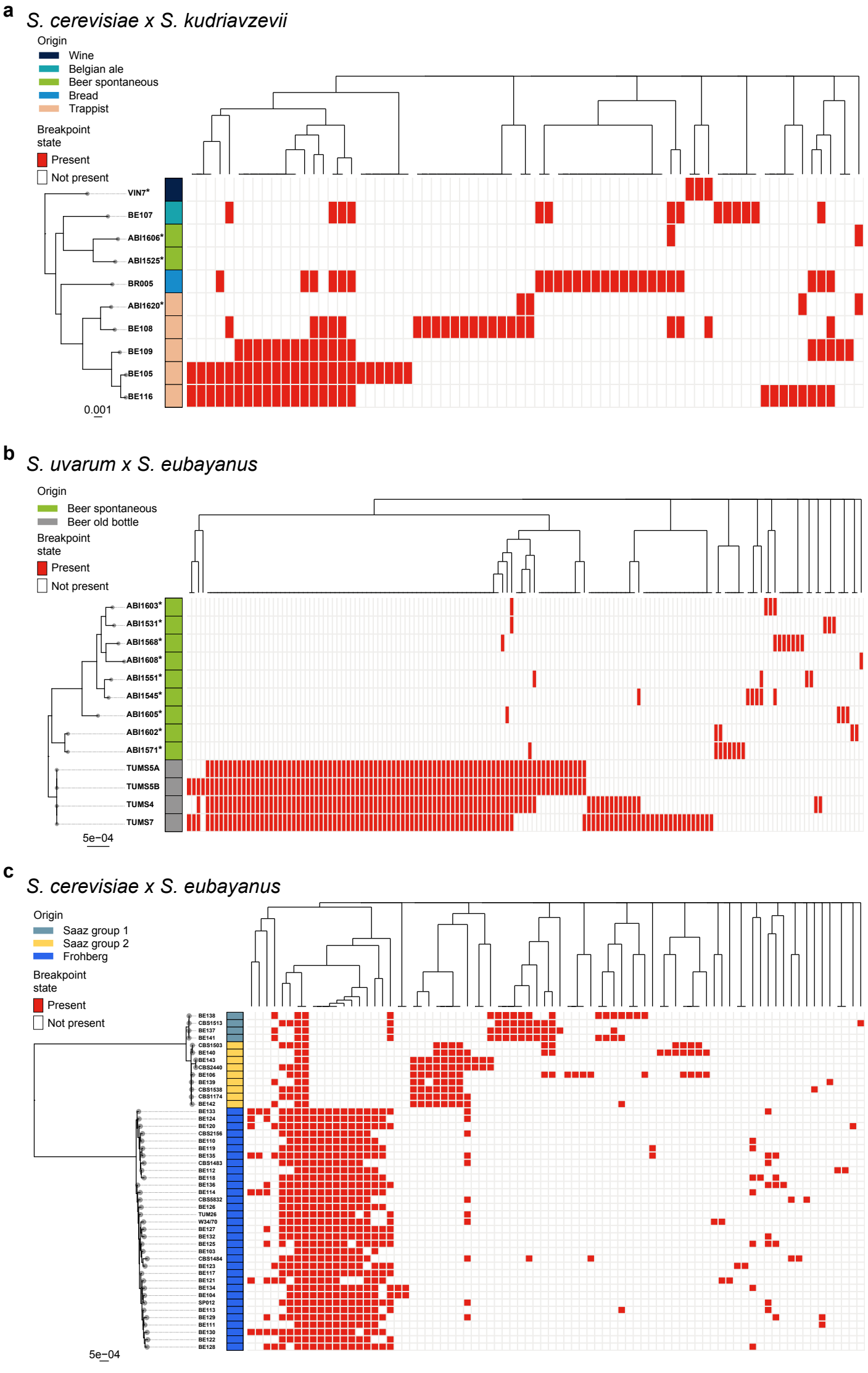
Lineage

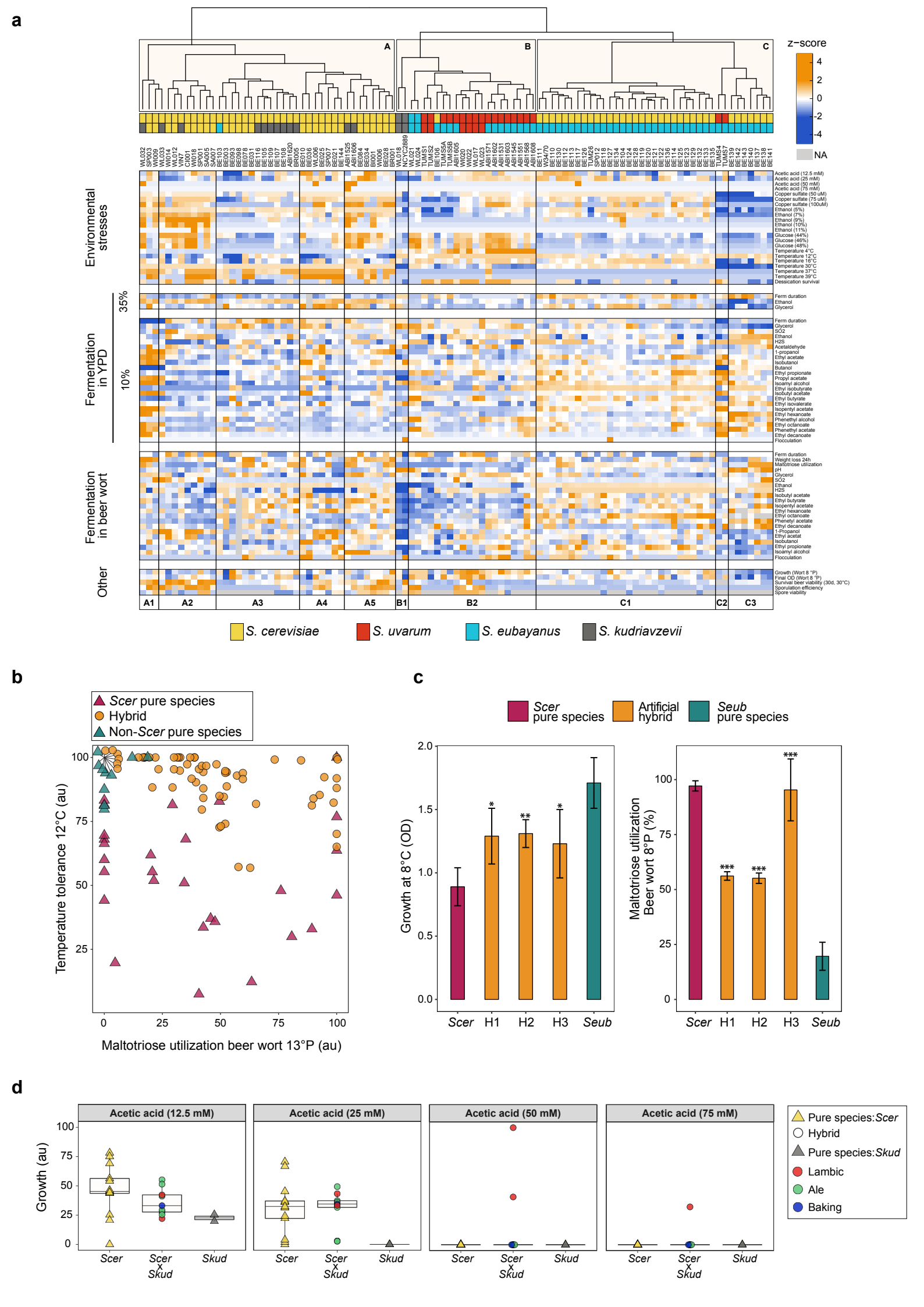
Beer 1						
1 Hefeweizen	5 Mixed	9 Laboratory	13 Asia	17 South America A	21 Asia wild
2 Belgium/Germany	6 Wine	10 Kombucha	14 Patagonia B1	18 South America B	
3 British	7 Beer2	11 West-Africa	15 Admixture	19 Australasia	
4 US	8 Mediterranean Oak	12 North-America/Japan	16 Patagonia A 1-2	20 Europe wild	

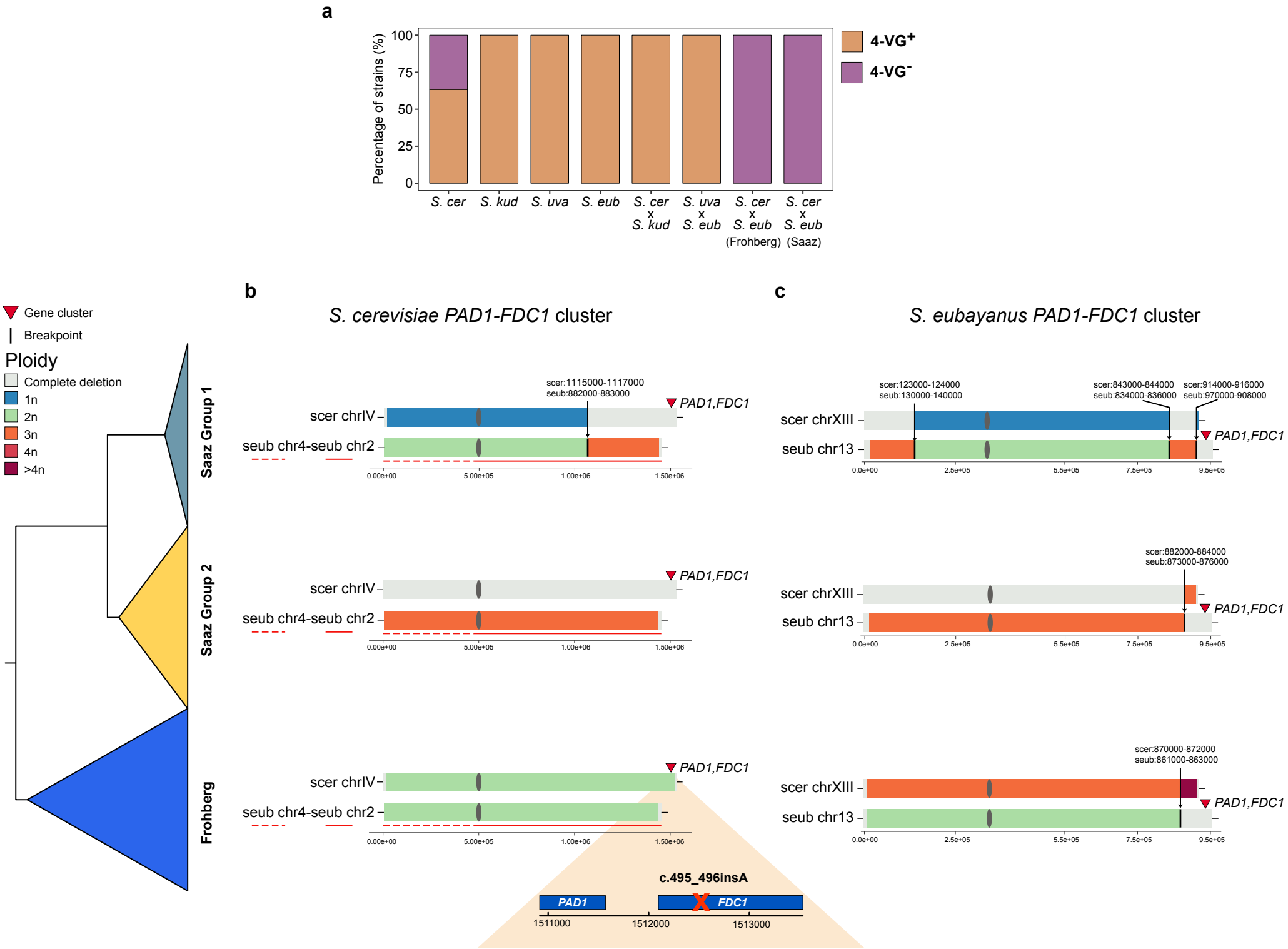
Interspecific hybrid type

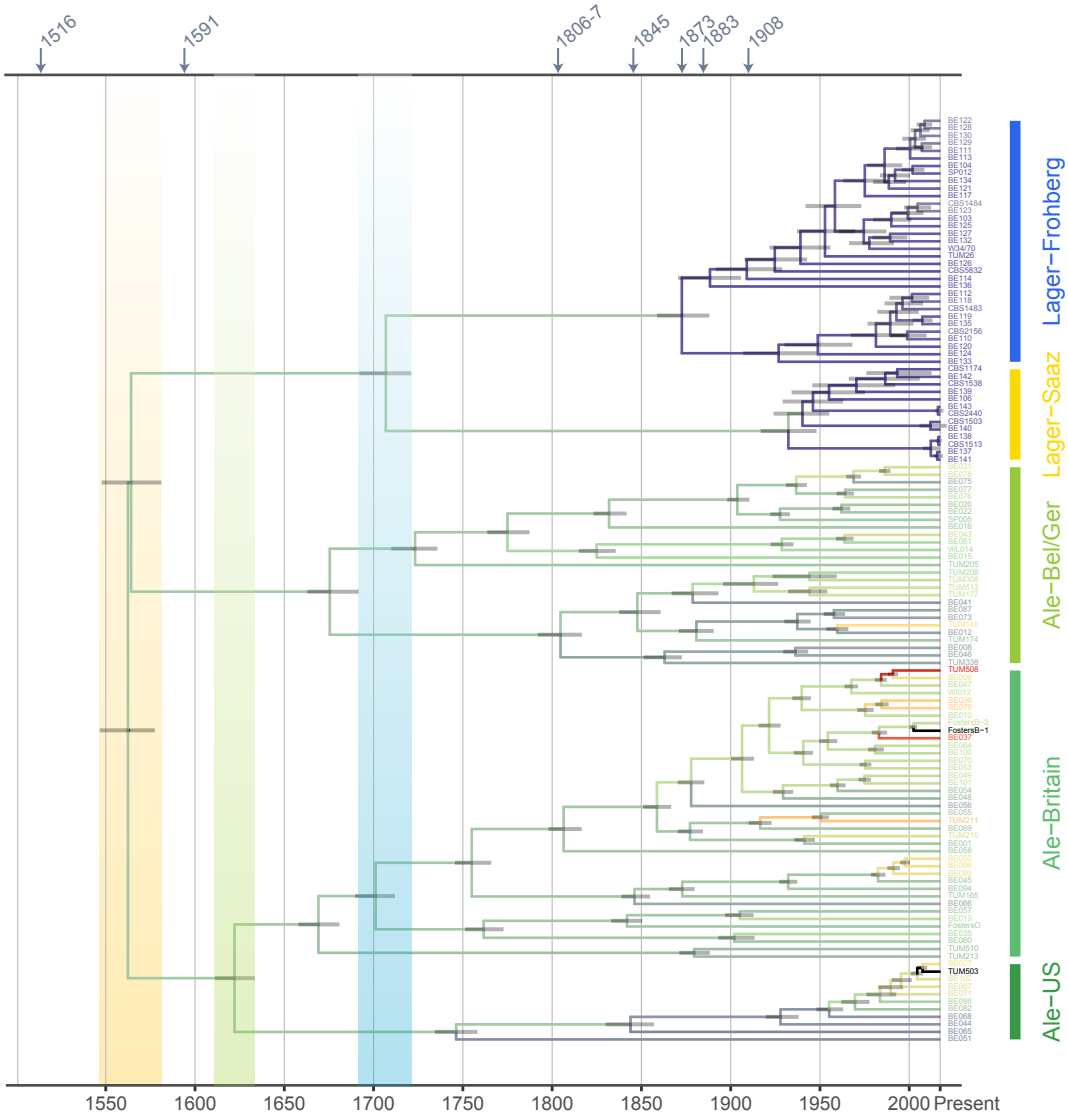
S. cerevisiae x S. eubayanus (Frohberg) *S. cerevisiae x S. eubayanus* (Saaz) *S. cerevisiae x S. kudriavzevii* *S. eubayanus x S. uvarum*


Calculated Ploidy
0
1n
2n
3n
4n
>4n
a
S. cerevisiae
S. kudriavzevii
Density all strains
Ploidy
Hybrid BE109
Ploidy S.cer
Ploidy S.kud
b
S. uvarum
S. eubayanus
Density all strains
Ploidy
Hybrid TUMS4
Ploidy S.uva
Ploidy S.eub
c
S. cerevisiae
S. eubayanus
Density all strains
Ploidy
Hybrid BE140
Ploidy S.cer
Ploidy S.eub
d
S. cerevisiae
S. eubayanus
Density all strains
Ploidy
Hybrid BE104
Ploidy S.cer
Ploidy S.eub


a
S. cerevisiae x S. kudriavzevii
Origin
Wine
Belgian ale
Beer spontaneous
Bread
Trappist
Breakpoint state
Present
Not present
VIN7*
BE107
ABI1606*
ABI1525*
BR005
ABI1620*
BE108
BE109
BE105
BE116
0.001
b
S. uvarum x S. eubayanus
Origin
Beer spontaneous
Beer old bottle
Breakpoint state
Present
Not present
ABI1603*
ABI1531*
ABI1568*
ABI1608*
ABI1551*
ABI1545*
ABI1605*
ABI1602*
ABI1571*
TUMS5A
TUMS5B
TUMS4
TUMS7
5e−04
c
S. cerevisiae x S. eubayanus
Origin
Saaz group 1
Saaz group 2
Frohberg
Breakpoint state
Present
Not present
BE138
CBS1513
BE137
BE141
CBS1503
BE140
BE143
CBS2440
BE106
BE139
CBS1538
CBS1174
BE142
BE133
BE124
BE120
CBS2156
BE110
BE119
BE135
CBS1483
BE112
BE118
BE136
BE114
CBS5832
BE126
TUM26
W3470
BE127
BE132
BE125
BE103
CBS1484
BE123
BE117
BE121
BE134
BE104
SP012
BE113
BE129
BE111
BE130
BE122
BE128
5e−04

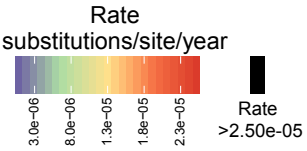

a
A
B
C
z-score
4
2
0
-2
-4
NA
Environmental stresses
Fermentation in YPD
35%
10%
Fermentation in beer wort
Other
A1
A2
A3
A4
A5
B1
B2
C1
C2
C3
S. cerevisiae
S. uvarum
S. eubayanus
S. kudriavzevii
b
Scer pure species
Hybrid
Non-Scer pure species
Temperature tolerance 12°C (au)
Maltotriose utilization beer wort 13°P (au)
c
Scer pure species
Artificial hybrid
Seub pure species
Growth at 8°C (OD)
Maltotriose utilization Beer wort 8°P (%)
Scer
H1
H2
H3
Seub
d
Acetic acid (12.5 mM)
Acetic acid (25 mM)
Acetic acid (50 mM)
Acetic acid (75 mM)
Growth (au)
Scer
Scer x Skud
Skud
Pure species:Scer
Hybrid
Pure species:Skud
Lambic
Ale
Baking


a
Percentage of strains (%)
100
75
50
25
0
4-VG+
4-VG-
S. cer
S. kud
S. uva
S. eub
S. cer x S. kud
S. uva x S. eub
S. cer x S. eub (Frohberg)
S. cer x S. eub (Saaz)
Gene cluster
Breakpoint
Ploidy
Complete deletion
1n
2n
3n
4n
>4n
Saaz Group 1
Saaz Group 2
Frohberg
b
S. cerevisiae PAD1-FDC1 cluster
scer:1115000-1117000
seub:882000-883000
PAD1,FDC1
scer chrIV
seub chr4-seub chr2
0.00e+00
5.00e+05
1.00e+06
1.50e+06
c
S. eubayanus PAD1-FDC1 cluster
scer:123000-124000
seub:130000-140000
scer:843000-844000
seub:834000-836000
scer:914000-916000
seub:970000-908000
scer chrXIII
seub chr13
0.0e+00
2.5e+05
5.5e+05
7.5e+05
9.5e+05
scer:882000-884000
seub:873000-876000
scer:870000-872000
seub:861000-863000
c.495_496insA
PAD1
FDC1
1511000
1512000
1513000



Historical data

1516	Bavarian Beer Purity Law
1591	Munich Hofbräuhaus established in Bavaria
1806	Gabriel Sedlmayr the Elder becomes head brewer at the Munich Hofbräuhaus.
1807	Gabriel Sedlmayr the Elder acquires the Spaten brewery
1845	J.C. Jacobsen, Gabriel Sedlmayr's student, founds Carlsberg and carries Bavarian yeast to Copenhagen
1873	Mechanical refrigeration was invented and implemented in breweries
1883	First yeast pure culture isolation by E.C. Hansen at Carlsberg (Unterhefe Nr. 1)
1908	E.C. Hansen isolates Unterhefe Nr. 2 at Carlsberg



Supplementary Information

Interspecific hybridisation facilitates niche adaptation in beer yeast

Gallone B., Steensels J., Mertens S., Dzialo M.C., Gordon J.L., Wauters R., Theßeling F.A., Bellinazzo F., Saels V., Herrera-Malaver B., Prahl T., White C., Hutzler M., Meußdoerffer F., Malcorps P., Souffriau B., Daenen L., Baele G., Maere S.*, Verstrepen K.J*.

Contents

***Correspondence to:**

Kevin J. Verstrepen
kevin.verstrepen@kuleuven.vib.be

Steven Maere
steven.maere@ugent.vib.be

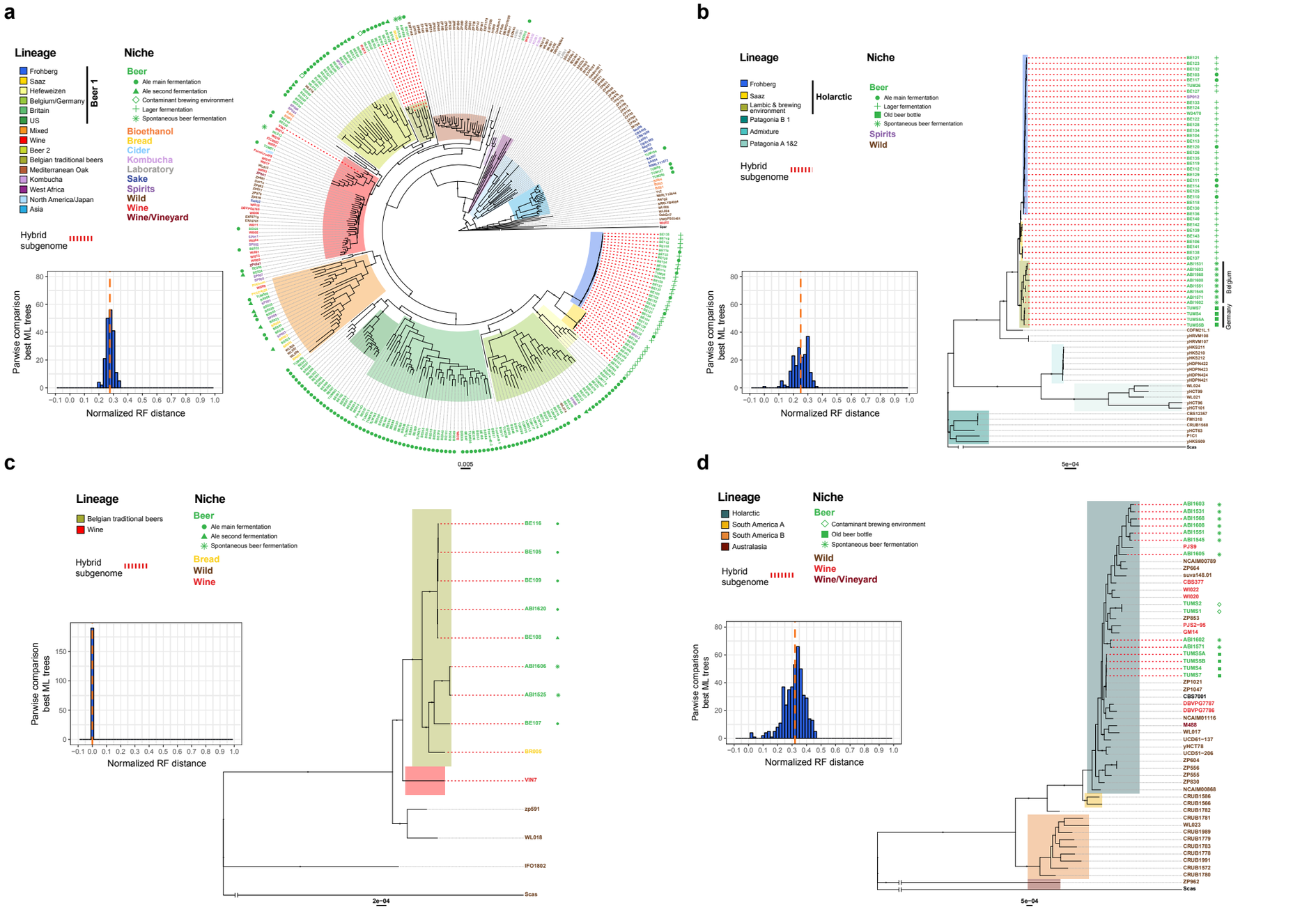


Figure S1. Expanded phylogenies of *Saccharomyces* genomes and subgenomes. Best maximum likelihood (ML) phylogenetic trees inferred for (a) *S. cerevisiae*, (b) *S. eubayanus*, (c) *S. kudriavzevii* and (d) *S. uvarum* (sub)genomes. Lineages are indicated by shaded boxes on the trees. Strain names are coloured according to isolation origin (niche) and beer strains are further divided into style categories (symbols). Hybrid subgenomes are highlighted by red dashed lines. Bar graphs in bottom left panels represent pairwise comparisons across all best -scoring ML trees expressed as Robinson-Foulds (RF) distances. Deep branches supported by more than 70% of bootstrap replicates are indicated by a grey dot.

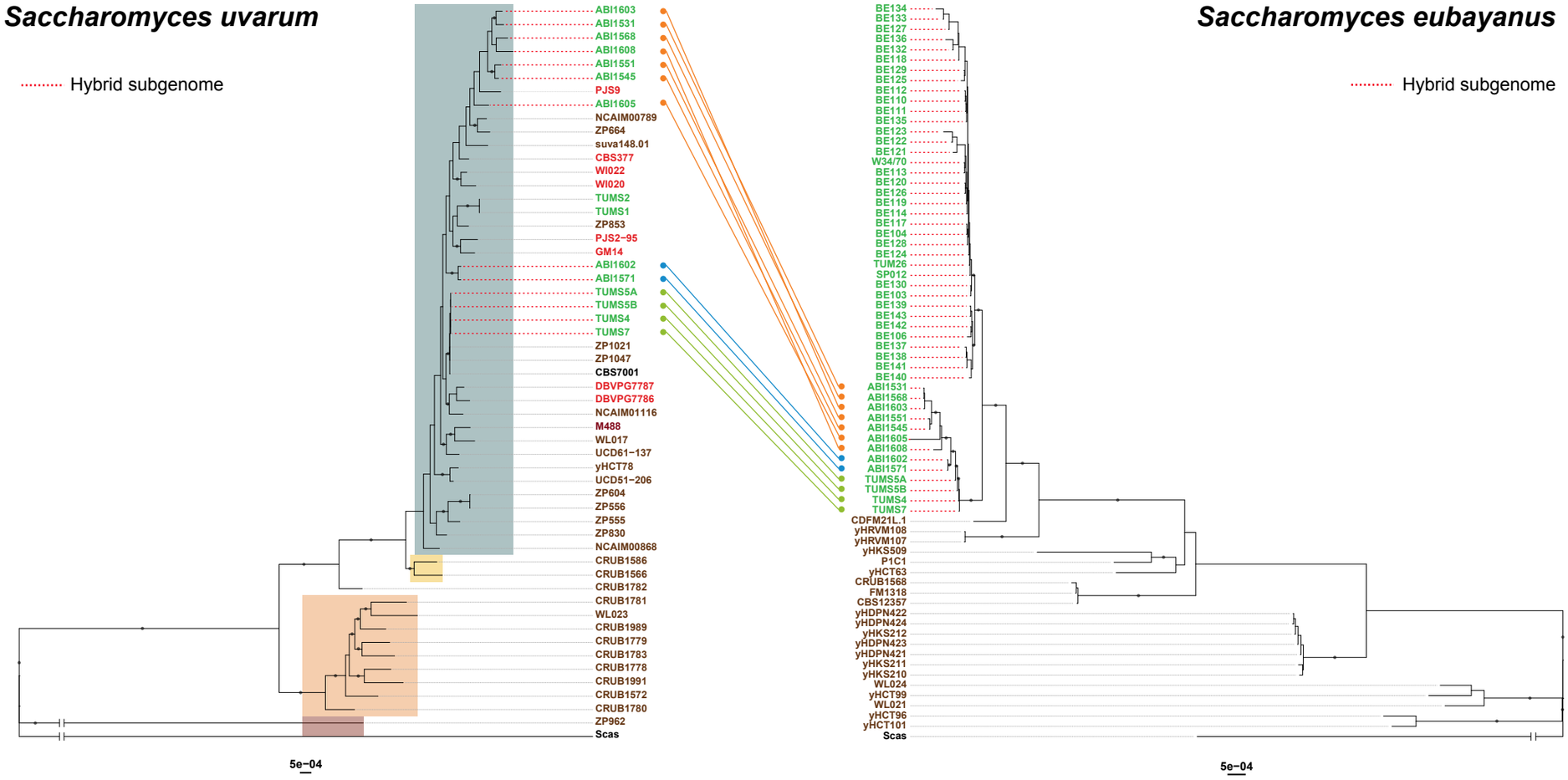


Figure S2. Tanglegram depicting the relationships between *S.uva* and *S.eub* pure species and interspecific hybrids. The maximum likelihood (ML) tree of *S.eub* is calculated on a reduced set of genes (166 single-copy orthologs) in order to include strain ABI1605, missing in the *S.eub* phylogenetic tree reported in Figure 1 and Figure S1b, due to the limited genomic contribution of *S.eub* in this strain. Each line of the tanglegram represents one *S.uva* x *S.eub* hybrid. Different line colours indicate different hybrid origins.

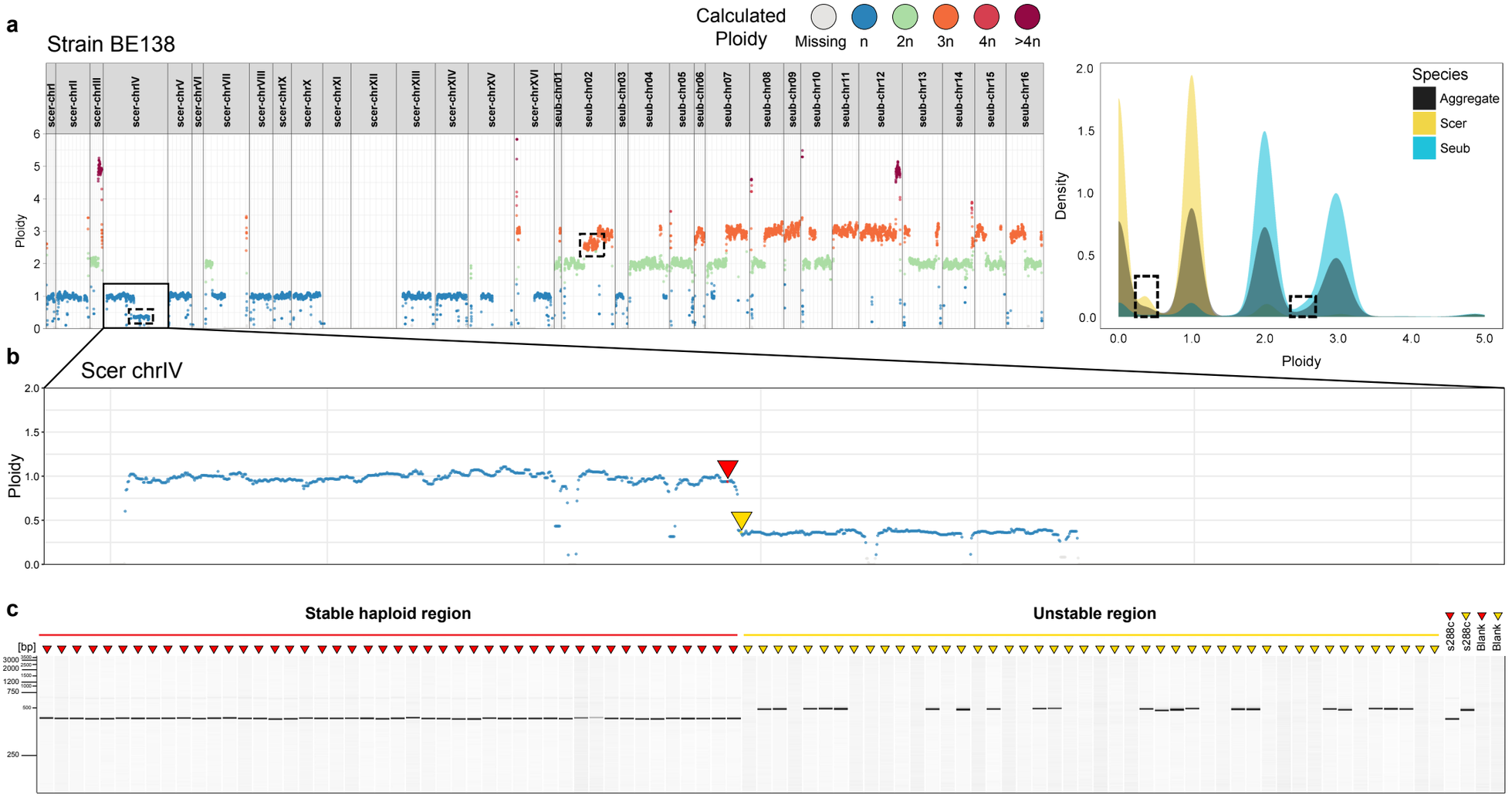


Figure S3. Genomic instability and copy-number heterogeneity in BE138. Detailed example of copy-number heterogeneity in a *S.cer* x *S.eub* hybrid, BE138. (a) Left panel: chromosomal level ploidy calculated in 1kb genomic regions based on read depth. Right panel: total density of calculated ploidy per subgenome, and aggregated for the full hybrid (colours). Black dashed squares highlight the presence of regions with an intermediate ploidy level relative to the calculated distribution. (b) Zoom-in on the ploidy profile of *S.cer* ChrIV in BE138 harbouring a region with intermediate ploidy level (between 0 and 1) compared to the haploid level calculated for the chromosome. Coloured triangles indicate the regions targeted by PCRs, red (stable) region, yellow (unstable, heterogeneous region). (c) Gel electrophoresis of the PCR products, demonstrating the presence/absence of the stable region (red triangles) and the unstable region (yellow triangle) across 45 random colonies of strain BE138.

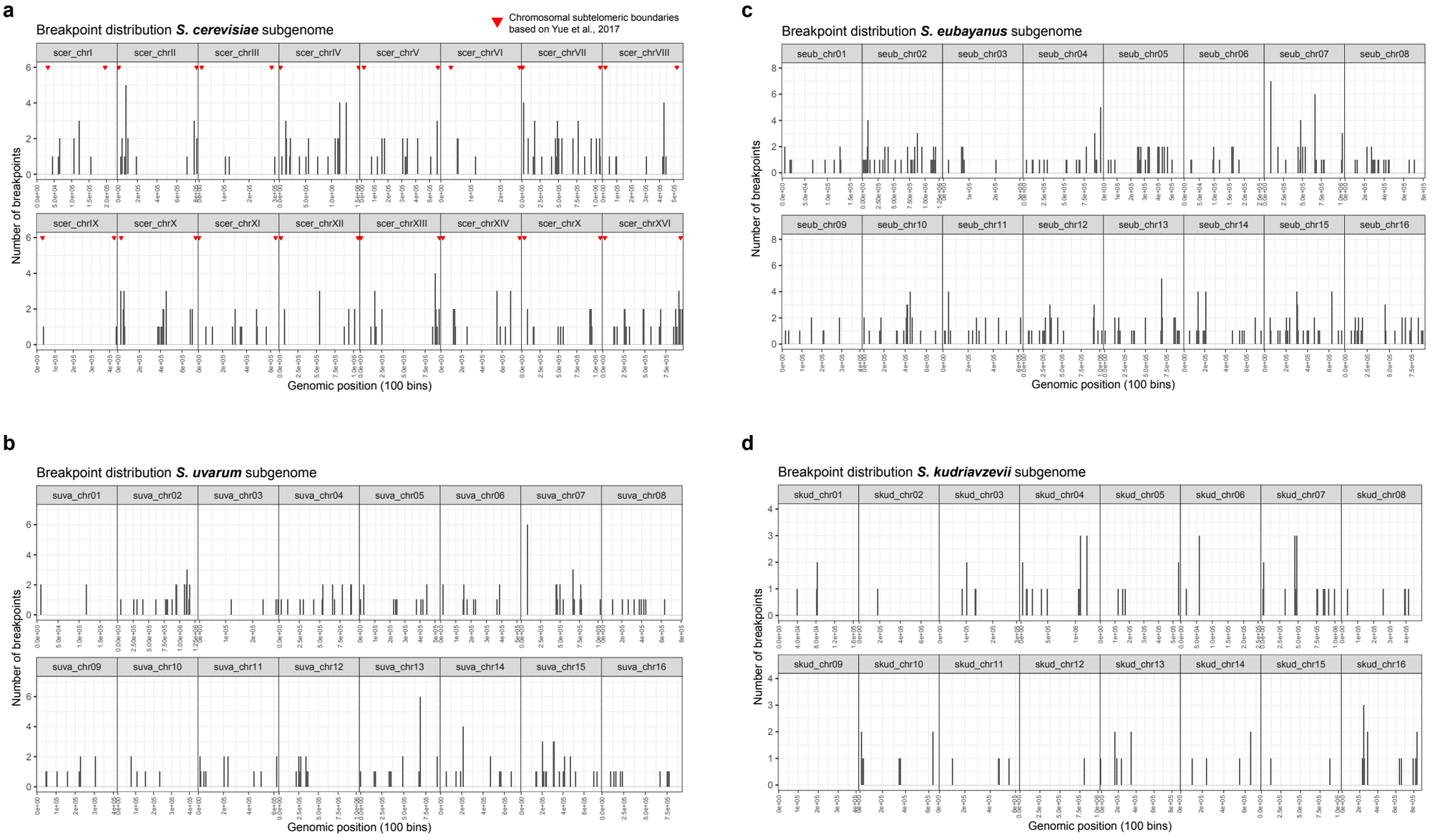


Figure S4. Whole-genome distribution of chimeric breakpoint windows in interspecific hybrids, aggregated accross subgenomes. Histogram depicting the distribution of chimeric breakpoint along the *S.cer* (a), *S.uva* (b), *S.eub* (c) and *S.kud* (d) subgenomes (see Supplementary Table 1 for a list of strains used in the breakpoint analysis). Raw data are reported in Supplementary Table 2.

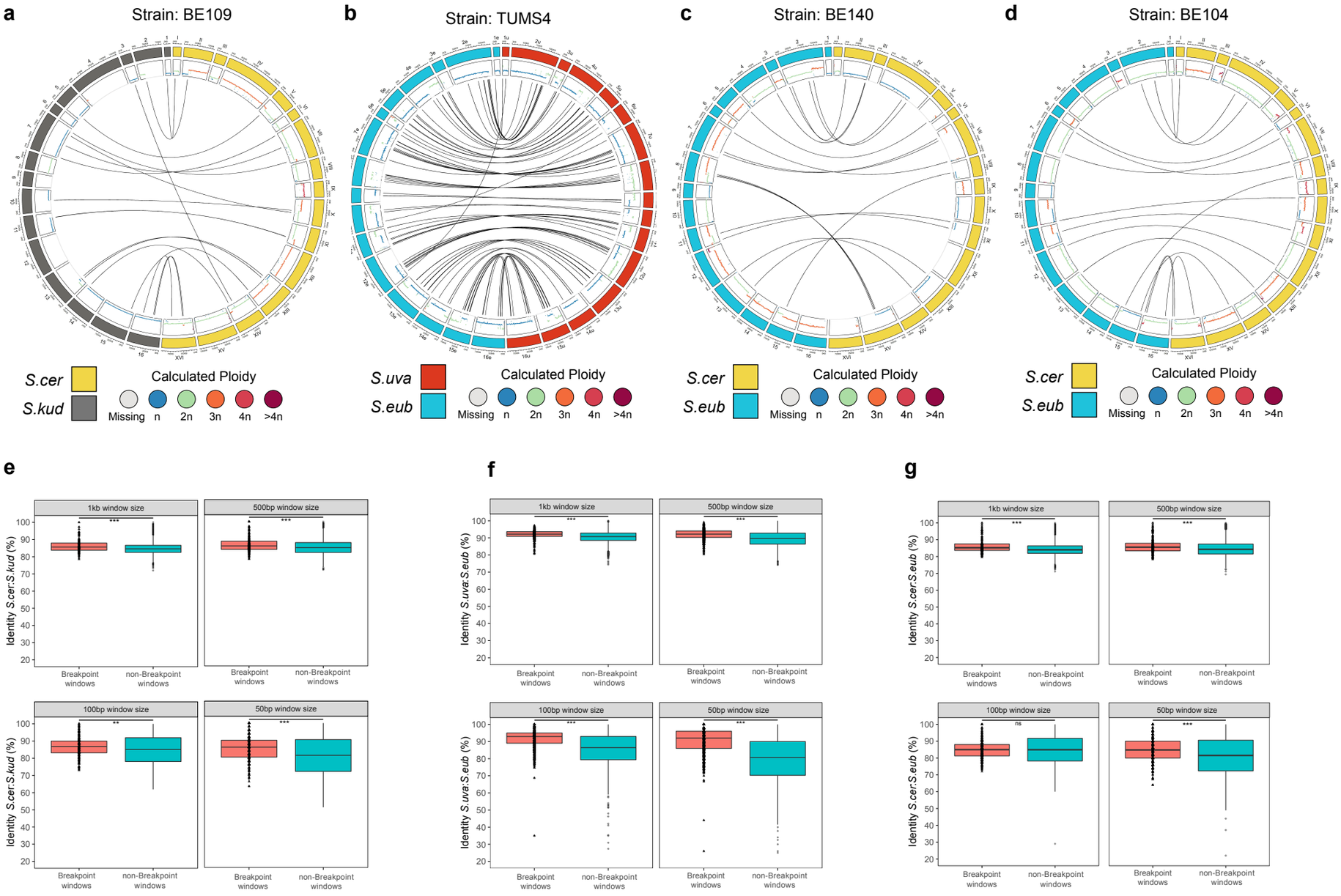


Figure S5. Detailed view of chimeric breakpoint windows in interspecific hybrids and nucleotide identity across subgenomes. Chimeric breakpoints identified in four representative hybrids for (a) *S.cer* x *S.kud* (BE109), (b) *S.uva* x *S.eub* (TUMS4), (c) *S.cer* x *S.eub* - Saaz (BE140) and (d) *S.cer* x *S.eub* - Frohberg (BE104). Outer tracks of circular plots represent reference chromosomes coloured according to species. Inner tracks represent chromosomal ploidy levels of the hybrid genome for each reference chromosome. Black lines connect chimeric breakpoint locations across the two subgenomes. (e, f, g) Nucleotide identity (%) between species (subgenomes) calculated in 1kb, 500b, 100bp and 50bp genomic windows within (e) *S.cer* x *S.kud*, (f) *S.uva* x *S.eub* and (g) *S.cer* x *S.eub* hybrids. Nucleotide identity (%) of windows that contain chimeric breakpoints (red) is compared to windows that do not contain chimeric breakpoints (blue) using the Wilcoxon Signed Rank Test and significance levels are indicated by asterisks: ns (P-value > 0.05),* (P-value ≤ 0.05), ** (P-value ≤ 0.01), and *** (P-value ≤ 0.0001).

6. Supplementary Note 1. A brief history of lager brewing and lager yeasts

In this supplementary note we describe major historical events in lager brewing and provide information on the origin of several key lager strains included in this study. The combination of historical information and the genetic data obtained in this study allows us to better understand lager yeast population structure, and enables us to identify key environmental drivers of their domestication.

First, we will provide a concise description of lager beer history, starting from its origin in 13th century Germany to its global spread in the late 19th century. Next, we trace back the origins of present-day lager yeast by providing historical information on the yeast and yeast treatment of some key lager breweries in old Bavaria and discuss the impact of pure yeast culturing. Lastly, we provide a historical trace of several yeasts included in this study.

Lager beer

'Lager' (March-) beers emerged in the 13th century. They are based on a brewing process whereby after a first fermentation, enough fermentable medium remains for a second, slower fermentation in barrels. From the 15th century on, a different variant of lager brewing was developed in Franconia and Bavaria (Germany), in which fermentation and storage were typically performed in cold cellars. This environment favoured so-called 'bottom-fermenting' yeasts, which accumulated at the bottom of the barrels, over 'top-fermenting' yeasts. Bavarian legislation subsequently restricted the production of traditional 'top-fermented' beers (Bavarian purity law, 1516), and therefore promoting bottom fermentation. In the 18th century, Munich brewers further optimized the production process of lager beer and began to build large storage (lagering) cellars. By the beginning of the 19th century, the Munich lager beer technology had maturated into a simple and robust process, which proved ideally suited for industrial beer production. Until the 19th century, this process was confined to the region which today comprises Bavaria, but in the first half of the 19th century, spread to all over the world. Interestingly, a similar brewing practice was also developed in Bohemia (westernmost region of Czech Republic), in parallel with the Bavarian brewers. However, a key difference between the two regions was that in Bohemia the yeast inoculum was taken always from the top layer (hence favouring top-fermenting yeasts), at least until the 19th century.

Instrumental to the popularization of the lager beer production process was a unique collaboration between brewers, particularly the Sedlmayr family, and scientists. In 1836, Professor Cajetan Kaiser of the Munich Technical University started a brewing course coupled to a practical training at Gabriel Sedlmayr's Spaten brewery (Spatenbräu) in Munich. This course was quite international, as only one third of the 900 brewers that participated in the course between 1836 and 1856 originated from Bavaria. In a promotion leaflet for the world exhibition at Paris in 1867, Gabriel Sedlmayr claims that over 300 of the contemporary top brewers had practiced at his brewery, amongst which pioneering

brewers such as Jacob Christian Jacobsen, the founder of the Carlsberg brewery in Copenhagen, Denmark. Because of these trainings, both knowhow and the bottom-fermenting yeast strains of Munich itself spread all over the world, being at that time the sole source for lager beer-producing yeasts. It is the progeny of these yeasts that are present in today's strain collections.

Lager yeast

In this section, we discuss how lager yeasts originated, how they were distributed amongst brewers and how they ended up in today's culture collections. Based on our genome analyses (phylogenetics and molecular dating) and historical information, we put forward some hypotheses on the evolutionary path of lager yeast.

Lager yeast origin

Using our large dataset of beer yeast genomes (both ale and lager), we performed an absolute dating analysis to put the genetic relatedness and the phylogenetic relationships of lager yeasts in a historical context. Vital for this analysis is the presence of a historical biogeographic event related to the Beer 1 clade, which serves as a calibration point. As previously reported, the split between British and US ale-producing strains can be traced back to the import of ale (and the yeast they carried) to North America by British settlers in the early 17th century. More specifically, the split between US and UK Beer 1 yeasts is thought to have occurred between 1607 and 1637. We used this calibration in a Bayesian molecular clock analysis to infer the age of evolutionary events related to the lager strain lineage and match the estimates with historical records. Still, it is important to stress that because of the paucity of calibration data available, our dating results can only be considered approximations.

Our analyses suggest that the most recent common ancestor of present-day ale and lager beer strains dates to the mid-16th century, between 1547-1578 AD (**Figure 6, yellow highlight**). The lager yeast lineage splits off from the Belgium/Germany ale-lineage within the same time frame. This is in accordance with the first observations of 'bottom-fermenting' lager yeast occurrence in the 16th century in Munich lager breweries, both in terms of timing and geographical location. Already the first books on brewing in the 16th century emphasised the difference between top-fermenting yeasts (yeasts that would float to the top of the fermentation tank towards the end of fermentation, like *S. cerevisiae*) and bottom-fermenting yeasts (which would sink to the bottom, like lager yeasts). Johannes Placitomus for example, a German physician and pedagogue who published one of the first books on brewing in 1551, discriminates between 'flos *cerevisiae*' (top-fermenting) and 'feces', '*cerevisiae* sedimentum' (bottom-fermenting).

In addition to the genetic data, historical data also seem to favour emergence of lager yeast in the 16th century. The combination of the establishment of the Bavarian purity law in 1506 (which encouraged

brewing in winter times) and the occurrence of a little ice age in Europe around the same time, likely made resistance to very cold temperatures an increasingly important prerequisite for yeast in lager beer production. This could have sparked the selection for hybrids between *S. cerevisiae* and *S. eubayanus* in these environments. In support of this hypothesis, historical data regarding a legal dispute between Munich brewers and their neighbouring bakers indicate that the yeast in lager beer production was changing in the early 16th century. A popular type of bottom-fermented beer, the so-called 'Bohemian' beer, was brewed in Bavaria from around 1481-1517. The Munich bakers, who traditionally acquired yeast from the brewery to ferment their dough, filed a complaint stating that the yeast they received since introduction of this new beer style showed altered properties, which made it unsuited for baking. It is possible that the yeasts with altered properties were early lager yeast hybrids, as it is well known that lager hybrids are unsuited for baking.

Lager yeast evolution from the 16th to the 19th century and the origin of the 'Frohberg' and 'Saaz' lineages

The newly formed lager yeast hybrids were continuously cultivated under lager beer conditions, gradually adapting to the unique environment in Munich breweries. As described in the section on 'Lager beer', the production process of lager beers in Bavaria further matured, leading to a tightly controlled process (that is similar to the current state-of-the-art) by the early 19th century.

While the number of different hybrid lineages might have been substantial in those days, the phylogenetic tree of lager yeasts shows that there are only two archetypes ('Frohberg' and 'Saaz') retained in the present-day lager yeast population. The terms 'Frohberg' and 'Saaz' were coined at the VLB at Berlin by Delbrück and Lindner at the turn of the 20th century, describing two pure type strains with a characteristic fermentation behaviour. They had been isolated from bottom-fermenting cultures obtained from the Frohberg's brewery at Grimma (Saxony) and the Saaz brewhouse (Bohemia). Our data suggest that the split between the Saaz and Frohberg type occurred somewhere in the 18th century. And while the exact event that led to the split between Saaz and Frohberg remains elusive, historical documentation on Bavarian lager-type beer production and influential brewers allows us to reconstruct the path of the ancestors of yeasts from both lineages.

An important event that might have had an impact on the evolution of both lineages is the foundation of the Munich Hofbräuhaus in 1591. This brewery was considered a model brewery in Bavaria as it consistently produced top quality beers for the ducal court. Its brewing technology was often adopted by Munich brewers and its yeast cultures were freely distributed in the region. In 1806, Gabriel Sedlmayr the Elder was appointed as brewmaster at the Hofbräuhaus. He left only one year later, in 1807, to acquire the Spaten brewery in Munich. There, he started to brew lager-type beer, presumably with the bottom-fermenting yeast culture obtained at Hofbräuhaus. Sedlmayr and later his son transformed the Spaten brewery into a model brewery for modern, industrialized lager beer production

and a hotspot for lager brewing education (see section on 'Lager beer'). Among the many trainees studying the new technology there, was Jacob Christian Jacobsen. After this training, Jacobson successfully established lager-brewing in the traditional brewery he had inherited in 1845 (which later evolved into the Carlsberg brewery). Jacobsen later acknowledged that he had carried the Spaten brewery yeast culture by stagecoach from Munich to Copenhagen, cooling his tin box at every stop. In 1877, inspired by Louis Pasteur's notes on hygiene in beer production, Jacobson established a modern laboratory, where in 1883 Emile Christian Hansen isolated the first pure yeast culture. According to both Hansen and Jacobsen, this isolate descended directly from the yeast derived from the Spatenbräu in 1845. It was dubbed 'Unterhefe nr. 1', a strain that is also included in our phylogenetic tree as part of the **Saaz** lineage (Figure 6 – strain CBS1513 and BE137 Table S1). Later, in 1908, Hansen isolated another strain which was referred to as 'Unterhefe nr. 2' from the culture descending from the 1845 Spatenbräu yeast culture (Figure 6 – strain CBS1503 and BE140 and Table S1).

Tracing back the origin of TUM 26, a Frohberg-type strain sequenced in our study and present in the TUM (Technical University of Munich) collection, sheds some light on the origin of the **Frohberg** lineage. TUM 26 was often used for the lager production at the Weihenstephaner Staatsbrauerei, especially before the introduction of TUM 34 (see further). TUM 26 was isolated between 1945 and 1956 from a yeast culture from the Sternburg brewery Lützschena-Leipzig (Eastern Germany), which is only a few kilometres away from Frohberg's brewery at Grimma. The Sternburg brewery Lützschena-Leipzig was founded in 1836-37 by Maximillian von Sternburg, who transformed it into a model Lager brewery. The "Bavarian Beer-Steam-Brewery Lützschena" as it was named originally was built and run as a blueprint of the Munich original, and a master brewer of the Augustiner brewery (Munich) was hired as a consultant. The master brewer in all likelihood brought the yeast culture from the Augustiner brewery Munich to the Sternburg Brewery Lützschena-Leipzig to start the first brewing batch. Together, this historical information also points towards a Bavarian, and more specifically Munich, origin of the Frohberg lineage.

Lager yeast evolution in the late 19th - 20th century

As the presence of several brewery environments would favour the maintenance of strain diversity, one would assume that the large number of lager breweries from the 15th to 19th century resulted in a burst of diversity of lager yeasts. However, the time-calibrated phylogeny (**Figure 6**) shows that after the split from ale strains in the mid-16th century, two long ancestral branches lead to the most recent common ancestor (MRCA) of present-day Frohberg strains (late 19th century) and the MRCA of Saaz strains (early 20th century). Only after that, strains within the two lineages started to slightly diversify. This is indicative of the presence of a strong bottleneck followed by slow local divergence at the turn of the 20th century. While there are probably several factors influencing the diversification profile of lager strains (see the result section of the manuscript), the single most important event was the

isolation of pure cultures in the late 19th century and the willingness of brewers to share well-performing cultures with colleagues during this period. For example, Jacobsen decided to give his yeast to many colleagues, amongst which the Munich breweries Spatenbräu and Leistbräu, the experimental governmental brewery at Weihenstephan and the Heineken brewery. This omnipresence of only a few yeast variants in the breweries likely resulted in the disappearance of a large portion of the natural biodiversity present at that time. After this period, brewing became more competitive, and the attitude of strain sharing changed in the course of the early 20th century. This way, lager yeast strains began to diverge allopatrically, leading to the biodiversity we observe today.

One particularly interesting observation in the time divergence analysis is the decrease in evolutionary rate in the lager lineage compared to the ale lineage (**Figure 6** – branch colour). This drop coincides with the onset of mechanical refrigeration, invented by Carl von Linde in 1873 while working for the Spaten brewery. The lower temperature of lager brewing restricts growth rates, and (more importantly) the implementation of cold storage occurred earlier in lager brewing than in ale brewing. While historical records are not present for all strains, many lager strains (for example W34/70, TUM 26, Unterhefe nr. 1 and Unterhefe nr. 2) were stored and maintained at cold temperatures after their isolation, which drastically slowed down their evolution. This was not common practice yet for ale yeasts, which therefore kept evolving within the beer medium at a higher rate.

Origin of Weihenstephan 34/70, a well-known present-day lager yeast

The source of most of our lager strains is unknown so we are unable to directly pinpoint which brewery the yeasts originated from. Moreover, the loss of German culture collections during World War II, the arbitrary naming of isolates and the periodical changes in yeast systematics, hamper the possibility of tracing back present-day strains to their original source. However, for some iconic strains, some historical information is retained. As described above, Hansen isolated the Saaz strains Unterhefe nr. 1 and nr. 2 at the Carlsberg brewery, but most likely these two strains originated at the Hofbräuhaus, and were later used at the Spaten brewery in Bavaria. While these strains were initially popular amongst lager brewers, the vast majority of breweries today use Frohberg yeasts for their fermentations. In this lineage, Weihenstephan 34/70 (W34/70) is considered as the benchmark yeast. Probably the main reason for its popularity is the work of Prof. Narziß, who published a dissertation on lager yeast performance in 1956. In this work, he compared the performance of various lager yeasts available at that time, and concluded that TUM 34 (the ancestor of W34/70) was a robust and efficient strain that produced beers with an excellent aroma profile. TUM 34 (known as the ‘Hasenhefe‘), originate from the Hasen-Bräu in Augsburg. The exact date at which this yeast was transferred at TUM is not completely clear, but it must have occurred between 1947 and 1955. After Prof. Narziß‘s dissertation, TUM 34 was used in more and more breweries, due to the consulting activities of the Technical University of Munich (TUM). However, more than a decade later, in 1970, TUM 34 was re-

isolated from the Bavarian State Brewery Weihenstephan, after it ran for over one decade in that specific brewing environment, and called W34/70 or TUM 34/70 (depending on strain collection and reference). Many of today's lager brewing yeasts are direct descendants of this strain.

References

Anonymus, 1845. Ueber das bayerische Bierbrauereiverfahren. Dinglers Polytechnisches Journal 96 (CXIII.): 478–482.

Anonymus, 1886. Über Gärung und Hefe. Denglers Polytechnisches Journal (6. Reihe, 9. Band) 259: 419-424.

Aubry, L., 1885. Ein Beitrag zur Klärung und Richtigstellung der Ansichten über reine Hefe. Zeitschrift für das gesammte Brauwesen 7: 133-139.

Balling, C. J. N., 1845. Die Gährungschemie wissenschaftlich begründet und in ihrer Anwendung auf die Weinbereitung, Bierbrauerei, Branntweinbrennerei und Hefenerzeugung practisch dargestellt. Zweiter Band. Verlag der J. G. Calve'schen Buchhandlung. Friedrich Tempsky, Prag.

Camenisch, C., K. M. Keller, M. Salvisberg, B. Amann, M. Bauch, S. Blumer, R. Brázdil, S. Brönnimann, U. Büntgen, B. M. S. Campbell, L. Fernández-Donado, D. Fleitmann, R. Glaser, F. González-Rouco, M. Grosjean, R. C. Hoffmann, H. Huhtamaa, F Joos, A. Kiss, O. Kotyza, F. Lehner, J. Luterbacher, N. Maughan, R. Neukom, T. Novy, K. Pribyl, C. C. Raible, D. Riemann, M. Schuh, P. Slavin, J. P. Werner, and O. Wetter, 2016. The 1430s: a cold period of extraordinary internal climate variability during the early Spörer Minimum with social and economic impacts in north-western and central Europe. Climate of the Past 12 (11): 2107–2126.

Estor, Johann Georg, 1757. Bürgerliche rechtsgelehrsamkeit der Teutschen. Bd. 1. Immanuel Weldige, Marburg. p. 603.

Euler, H., and P. Lindner, 1915. Chemie der Hefe und der alkoholischen Gärung. Akademische Verlagsgesellschaft, Berlin.

Feise, W., 1928. Das Brauwesen der Stadt Eimbeck. Gesellschaft für die Geschichte u. Bibliographie d. Brauwesens, Berlin.

Gallone, B., S. Mertens, S. Crauwels, B. Lievens, K. J. Verstrepen, and J. Steensels, 2017. Genomics and Evolution of Beer Yeasts. in N. A. Bokulich and C. W. Bamforth edts. Brewing Microbiology. Cur-rent Research, Omics and Microbial Ecology. Caister Academic Press, Norfolk (UK). pp. 145-177.

Gattinger, K., 2007. Bier und Landesherrschaft: Das Weißbiermonopol der Wittelsbacher unter Maximilian I. von Bayern. Karl M. Lipp Verlag, München.

Gibson, B. R., Storgårds E., Krogerus K., and Vidgren V., 2013. Comparative physiology and fermentation performance of Saaz and Frohberg lager yeast strains and the parental species *Saccharomyces eubayanus*. Yeast 30: 255–266.

Hansen, E. C., 1895. Untersuchungen aus der Praxis der Gärungsindustrie: Beiträge zur Lebensgeschichte der Mikroorganismen. Dritte, vermehrte und neubearbeitete Auflage. R. Oldenbourg, München und Leipzig. p. 18.

Heiß, P., 1853. Die Bierbrauerei mit besonderer Berücksichtigung der Dickmaischbrauerei. Im Selbstverlag, München.

Hermbstädt, S. F., 1814. Chemische Grundsätze der Kunst Bier zu brauen. Carl Friedrich Amelang, Berlin.

Huntemann, H., 1970. Bierproduktion und Bierverbrauch in Deutschland vom 15. bis zum Beginn des 19. Jahrhunderts. Dissertation an der der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Georg-August-Universität zu Göttingen.

Irsigler, F., 1996. "Ind machden alle lant beirs voll"- Zur Diffusion des Hopfenbierkonsums im westlichen Hanseraum. In: Nahrung und Tischkultur im Hanseraum (G. Wiegelmann und R-E. Mohrmann Hrsg.), Waxsmann Verlag GmbH, Münster, Deutschland. pp. 377-bis 385.

Kobrer, C., 1581. Gründliche und Nütze Beschreibung der Weinhawer und Bierbrewer-Practick und der ganzen Kellermeister-Kunst. Regenspurg.

Koch, A., 1894. Jahresbericht über die Fortschritte in der Lehre von den Gährungsorganismen, 5. Jahrgang. Harald Bruhn, Braunschweig.

von Kreittmayr, W. X. A., 1761. Anmerkungen über den Codicem Maximilianeum Bavaricum Civilem. Zweyter Theil., München. pp. 1465-1466.

Kromer, B., and M. Friedrich, 2007. Jahresringchronologien und Radiokohlenstoff. Geographische Rundschau 59 (4): 50-55.

Landsteiner, E., 1999. The crisis of wine production in late sixteenth-century Central Europe: Climatic causes and economic consequences. Climatic Change 43: 323-334.

Letzing, H., 1990. Das Weißbierprivileg Herzog Wilhelms IV von Bayern für Hans VI. von Degenberg. Jahrbuch der Gesellschaft für die Geschichte und Bibliographie des Brauwesens 1994/1995, Berlin.

Lindner, P., 1891. Über die Erkennung der Heferassen und ihre photographische Darstellung. Wochenschrift für Brauerei 8: 815-816.

Lindner, P., 1901. Mikroskopische Betriebskontrolle in den Gärungsgewerben. Dritte, neubearbeitete Auflage. Verlagsbuchhandlung Paul Parey, Berlin.

Ling, A. R., and D. R. Nanji, 1921. On the longevity of certain species of yeast. Proceedings of the Royal Society Lond. B 92: 355-357.

Lorenz, H., 1916. Quellen zur städtischen Verwaltungs-, Rechts- und Wirtschaftsgeschichte von Quedlinburg.

Lüers, H., und F. Weinfurter, 1931. Die Einführung der Hefereinzucht in den Brauereibetrieb. in: H. Schulze-Besse (Hrsg.), Die Hefereinzucht in der Entwicklungsgeschichte der Brauerei. Gesellschaft für Geschichte und Bibliographie des Brauwesens e.V., Berlin. pp. 23-105.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner, 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. Science, 303 (5663): 1499-1503.

Meussdoerffer F., Zarnkow M. 2016. Das Bier: Eine Geschichte von Hopfen und Malz, Publisher C.H.Beck, Munich (Germany), ISBN-13 978-3406666674

Neumann, C., 1735. Lectiones publicae von vier Subjectis Diäteticis, Nemlich von den in hiesigen Gegenden gewöhnlichen und durch menschliche Hülffe zustande gebrachten Vierer Getränken. Gottlob Benjamin Fromman, Buchhändler des Waisenhauses in Züllichau.

Nguyen, H. V and T. Boekhout, 2017. 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 research, 17 (2): fox014.

Noback, G., 1874. Bier, Malz, so wie Maschinen und Apparate für Brauereien und Mälzereien. (Gruppe IV, Section 3). in: (C. T. Richter, edt.) Officieller Ausstellungs-Bericht herausgegeben durch die General-Direction der Weltausstellung 1873. K. K. Hof- und Staatsdruckerei, Wien.

Peterka, O., 1917. Die bürgerlichen Braugerechtigkeiten in Böhmen: eine rechtsgeschichtliche Untersuchung. J. G. Calve, Prag. pp. 125–127.

Placotomus, J., 1551. De natura et viribus cerevisiarum et mulsarum opusculum. Wittenberg.

Powell, C.D., Quain D.E, and Smart K.A., 2004. The impact of sedimentation on cone yeast heterogeneity. J. Am. Soc. Brew. Chem. 62: 8–17.

Reinke, O., 1891. Über den mit verschiedenen Heferassen zu erzielenden Vergährungsgrad, seine Bedeutung für die Würzeuntersuchung und die Bierbereitung. Wochenschrift für Brauerei VIII (39): 809-815.

Schönfeld, F., 1938. Obergärige Biere und ihre Herstellung. Zweite, vollständig neubearbeitete Auflage. Paul Parey, Berlin.

Sedlmayr F., 1951. Die Geschichte der Spatenbrauerei unter Gabriel Sedlmayr dem Älteren und dem Jüngeren 1807-1874. 2. Band. Hans Carl Verlag, Nürnberg. p. 189.

Seifert, J. A., 1818. Das Bamberger Bier, oder praktische, auf chemische Grundsätze gestützte Verfahrungsweise, Handgriffe und Gewerbs-Vortheile beim Brauen des Bamberger Bieres. Eigenverlag, Bamberg.

Sprotte, J., 2016. Die Geschichte der Lederer-Bräu zu Nürnberg. Jahrbuch der Gesellschaft für die Geschichte des Brauwesens, Berlin. pp. 149-181.

Teich, M., 2001: Bier, Wissenschaft und Wirtschaft in Deutschland 1800 - 1914. Ein Beitrag zur deutschen Industrialisierungsgeschichte, Publisher Böhlau, Vienna (Austria), ISBN-13 978-3205992394

Unger, R., 1998. Dutch brewing in the ninteenth century. in: (R. G. Wilson and T. R. Gourvish edts.) The dynamics of the international brewing industry since 1800. Routledge, London. pp.13-31.

Walther, A., A. Hesselbart, and J. Wendland, 2014. Genome sequence of Saccharomyces carlsbergensis, the world's first pure culture lager yeast. G3: Genes, Genomes, Genetics, 4 (5): 783-793.

Woltmann. K. L., 1827. Geschichte von Böhmen. 8. Band. August Rücker, Berlin. p. 306-311.

7. Supplementary Dataset 1. List of PCR primers used in this study

Primers for BR005	
Stable	Sequence
FW	AGAGGGAACACGAGTAATTGACA
RV	CACCAGCAAAATTAAAACCAGAG
Unstable	Sequence
FW	GGACGTTATGCTTTGCTGTTATC
RV	CAAAATTCTTCCCATAGTGATCG

Primers for ABI1620	
Stable	Sequence
FW	GGCGGGTCATATTCGTGTTA
RV	GGCGGGTCATATTCGTGTTA
Unstable	Sequence
FW	TGGTCGGAAAAGGATGAAAG
RV	GAAAGCAAAAATCCGTCTTCA

Primers for BE137	
Stable	Sequence
FW	AGTTGACCTAGACGAGCTGTTTG
RV	ACCTCTACAGCCTCACAACCATA
Unstable	Sequence
FW	TTTCGGCAATTCATGATAGAGAT
RV	GCTGATCTGTTCCATTCGTTATC

Primers for BE138	
Stable	Sequence
FW	CATGGGGATAAAGTGGTTCATAA
RV	TACTATGAAAAGGGCTGGATGAA
Unstable	Sequence
FW	AGGAAACATCAAACTGGAGATCA
RV	CTTCCAATGCAAGAACAGAATTT

8. Supplementary Table 1. Detailed information on strains included in this study.

[see the excel file]

9. Supplementary Table 2. Overview of chimeric breakpoint windows identified in each interspecific hybrid (rows) divided by hybrid-type (sheets). For read-pairs mapping on two different species the following information are reported: mapping chromosome on species 1, start coordinates of the read on species 1, end coordinates of the read on species 1, mapping chromosome species 2, start coordinates of the read on species 2, end coordinates of the read on species 2, breakpoint name (based on *S.cer* location for *S.cer* x *S.kud* and *S.cer* x *S.eub* hybrids and on *S.uva* location for *S.uva* x *S.eub* hybrids), mapping chromosome species 1 (1 kb window), start of the 1kb window in species 1, end of the 1kb window in species 1, mapping chromosome in species 2, start of the 1kb window in species 2, end of the 1kb window in species 2, read name, number of mismatches, length of read1, length of read2, summed length read1 and read2, normalised nucleotide identity (%) to the mapping location, annotated feature in species 1 at that location, annotation source for species 1, type of annotated feature in species 1, start coordinate annotated feature in species 1, end coordinate annotated feature in species 1, gene name species 1 if applicable, annotated feature in species 2 at that location, annotation source for species 2, type of annotated feature in species 2, start coordinate of annotated feature in species 2, end coordinate of annotated feature in species 2, gene name species 2 if applicable.

[see the excel file]

10. Supplementary Table 3. Phenotypic variation of *Saccharomyces* interspecific hybrids and *Saccharomyces* pure species for 76 traits depicted as Z-scores.

[see the excel file]