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 Prahl⁶, Christopher White⁶, Mathias Hutzler⁷, Franz Meußdoerffer⁷, Philippe Malcorps⁸, Ben Souffriau⁸, Luk Daenen⁸, Guy Baele⁹, 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

25 **Abstract**

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

38

39 **Introduction**

40 A canonical view of Darwinian evolution asserts that new species arise during extended
41 periods of gradual evolution and selection in combination with reproductive isolation.
42 However, amendments to Darwin's original theory have been proposed to incorporate other
43 mechanisms for evolution and speciation, including the emergence of new species through
44 interspecific hybridisation¹⁻³. Hybridisation provides a way to rapidly combine distinct
45 phenotypic features from established populations that converge in an unfamiliar ecological
46 niche. Sometimes, a hybrid's unique combination of phenotypes can enable it to thrive in the
47 new environment and outcompete its parental species⁴⁻⁶.

48 Examples of rapid niche specialisation via interspecific hybridisation are found across all
49 kingdoms of life. A new hybrid lineage of Darwin's finches with competitive beak
50 morphology emerged in only three generations⁷. Hybridisation of the sunflowers *Helianthus*
51 *annuus* and *Heliantus petiolaris* gave rise to three novel species capable of colonizing
52 previously untapped environments⁸⁻¹⁰. The emerging opportunistic fungal pathogen *Candida*
53 *metapsilosis* arose from a single hybridisation event between two non-pathogenic parental
54 species¹¹.

55 The *Saccharomyces* species complex provides numerous examples of interspecific
56 hybridisation. Despite the high sequence divergence between the species within this genus
57 (up to 20%¹²), prezygotic barriers are weak and species can successfully crossbreed^{13,14}. The
58 resulting hybrids are typically infertile, yet viable, and can reproduce asexually by budding¹⁴⁻
59 ¹⁷. Wild *Saccharomyces* interspecific hybrids are occasionally encountered^{18,19}, but the best-
60 known example, *Saccharomyces pastorianus*, was isolated from an industrial environment²⁰⁻
61 ²³.

62 Humans have historically utilised the capacity of *Saccharomyces* to convert high
63 concentrations of sugars into ethanol and carbon dioxide to produce a wide variety of
64 fermented products²⁴⁻²⁶. This long-term domestication process has resulted in hundreds of
65 different industrial strains, each with characteristics suitable for specific industrial
66 processes²⁷⁻³⁶. Strains used in the production of lager (pilsner) beer generally belong to the
67 species *S. pastorianus*, a hybrid of *S. cerevisiae* and *Saccharomyces eubayanus*^{20-23,37,38}. Two
68 archetypical *S. pastorianus* sublineages exist, dubbed 'Frohberg' and 'Saaz', each with its
69 own distinct fermentation properties³⁹. Several other industrial interspecific hybrids have
70 been isolated, including *S. cerevisiae* x *Saccharomyces uvarum* x *S. eubayanus* triple hybrids

71 in wine and cider, and *S. cerevisiae* x *Saccharomyces kudriavzevii* hybrids in ale beers and
72 wine fermentation vessels^{40–47}.

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

87 **Results**

88 **Origins and diversity of *Saccharomyces* interspecific beer hybrids**

89 For several years we have been collecting and sequencing yeasts isolated from different
90 industrial niches, including beer, wine, bread, sake, chocolate and liquor fermentations. As
91 expected, the majority of isolates were *S. cerevisiae*, the yeast species most commonly
92 associated with the production of fermented foods and beverages^{28,48,49}. However, our
93 analysis revealed that a surprisingly large fraction (~25%) of sequenced isolates were
94 interspecific hybrids. While some were isolated from lager beer fermentations, known to be
95 driven by hybrid yeasts, many were collected from other beer niches, such as Trappist beers,
96 spontaneously fermented ‘Lambic’ beers, and old beer bottles or equipment (**Table S1**). We
97 identified three distinct species compositions across the hybrids investigated: *S. cerevisiae* x
98 *S. eubayanus* (*S.cer* x *S.eub*), *S. cerevisiae* x *S. kudriavzevii* (*S.cer* x *S.kud*) and *S. eubayanus*
99 x *S. uvarum* (*S.eub* x *S.uva*). Phylogenetic trees were built to position the hybrid subgenomes
100 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 ($\pi=1.18\text{E-}03$ and $\pi=3.65\text{E-}04$, respectively). The *S.eub* subgenomes of the *S.cer* x *S.eub* hybrids similarly form a monophyletic group within the previously identified Holarctic clade⁵² (**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 Tibet⁵⁴. 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 difficult⁵².

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 bakers⁵⁵ (**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

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

144 The *S.eub* subgenomes of the *S.eub* $\times$ *S.uva* hybrids form a monophyletic sister clade to those
145 of the *S.cer* $\times$ *S.eub* lager strains, indicating that the *S.eub* parents were closely related. In
146 contrast to the monophyly of the *S.eub* subgenomes, the *S.uva* subgenomes are genetically
147 clearly separated, indicating that at least three hybridisation events gave rise to the *S.eub* $\times$
148 *S.uva* hybrids (**Figure S1D** and **Figure S2**). The *S.eub* subgenomes clearly separate
149 according to geographical isolation (Germany vs. Belgium), suggesting that a single or few
150 closely related *S.eub* strains formed multiple hybrids, which then evolved and diverged
151 independently. The genetic diversity across the *S.uva* subgenomes is significantly higher than
152 across the corresponding *S.eub* subgenomes (average nucleotide diversity $\pi=2.11\text{E-}03$ and
153 $\pi=4.66\text{E-}04$ respectively, one-sided Mann–Whitney U test, $P < 2.2 \times 10^{-16}$) and there is no
154 clear niche substructure, suggesting that *S.uva* strains move more freely across environments.
155 Indeed, in contrast to *S.eub*, *S.uva* has been isolated all over the world from a wide array of
156 niches, natural and man-made⁶⁰. Moreover, the species is a known contaminant in brewing
157 environments^{37,61,62}. Coupled with the fact that all of the *S.eub* $\times$ *S.uva* hybrids were isolated
158 from spontaneous fermentations, old bottles or brewing equipment, it is likely that the
159 hybridisations between *S.eub* strains and *S.uva* contaminants occurred within the brewing
160 environment. All *S.uva* hybrid subgenomes are also closely related to European strains
161 (**Table S1**), and thus the hybrids likely originated in Europe. However, as the population
162 structure of *S.eub* and *S.uva* is not described as elaborately as that of *S.cer*, an Asian origin
163 (as suggested in²³) 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 aneuploidies 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* $\times$ *S.kud* hybrids plus one *S.cer* $\times$ *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* $\times$ *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* $\times$ *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 process⁷³. 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 (*PADI*) and ferulic acid decarboxylase (*FDCI*)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 *PADI* and/or *FDCI*^{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* *PADI*-*FDCI* 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 *PADI*-*FDCI* 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 yeasts²⁸, 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 AD⁷⁶. 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

strains⁷⁷. The first isolated culture ('Unterhefe nr. 1') was disseminated across many breweries in central Europe, and rapidly implemented in their production process⁷⁸. 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* $\times$ *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 $\pm$ 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 *PADI-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 MasterPure™ 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-authors⁸¹. The starting set of genes was reduced to 4,125 genes after including *S.eubayanus* FM1318⁵⁶ 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 criterion⁹⁶ (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 $\geq 99\%$ 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 $\leq 3.00 = 99.8\%$; 200 replicates for the “*S.eub*” phylogeny: average WRF = 1.40%, percentage of permutations in which the WRF was $\leq 3.00 = 99.7\%$; 600 replicates for the “*S.uva*” phylogeny: average WRF = 1.36%, percentage of permutations in which the WRF was $\leq 3.00 = 99.2\%$; 50 replicates for the “*S.kud*” phylogeny: average WRF = 0.26%, percentage of permutations in which the WRF was $\leq 3.00 = 100\%$. 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 fastStructure¹⁰⁰ 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 events²⁸. We assessed the performance of several molecular clock models using BEAST (v1.10)⁹⁹ in combination with BEAGLE 2.1.2¹⁰¹. Specifically, we analysed the data using a strict clock, an uncorrelated relaxed clock with an underlying lognormal distribution¹⁰², a random local clock¹⁰³ 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 1.7¹⁰⁵. 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 sampling¹⁰⁶. 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 package⁹⁹.

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->

666 project.org/package=caTools) in R. Second, the density of the smoothed coverage was plotted
 667 per subgenome and for the full hybrid genome. In order to identify the mean value of
 668 smoothed coverage corresponding to the haploid ploidy level, a Gaussian mixture model was
 669 fitted to the density distribution and mean and standard deviation were calculated for each
 670 peak (peaks are a proxy for ploidy levels detected in the hybrid genome). Third, the smoothed
 671 raw coverage values were divided by the mean value of the haploid peak to obtain estimated
 672 ploidy values for each pre-computed window in the genome. Due to the presence of coverage
 673 depth noise and potential copy number heterogeneity in the population of sequenced cells,
 674 “in-between” (non-integer) ploidy levels were detected. In order to define integer ploidies
 675 and detect ploidy shifts, a second Gaussian mixture model was fitted on the distribution of
 676 estimated ploidy values for each strain. This allowed the identification of mean and standard
 677 deviation of each ploidy peak and the definition of “ploidy shift boundaries” based on the
 678 intersection points between the ploidy modes in the mixture model. The identification of
 679 breakpoint windows and their ploidy context was then followed by the identification of the
 680 exact mapping location of the reads within the breakpoint window across the two
 681 subgenomes for each hybrid strain based on the following steps: i) extract the reads from the
 682 selected breakpoint windows; ii) retain reads that are still paired after the identification and
 683 selection of breakpoint windows; iii) calculate for each pair their summed edit distance from
 684 the corresponding reference sequence, normalise it by the summed length of the two reads;
 685 iv) select reads with a percentage identity $\geq 95\%$ against the corresponding reference
 686 location; iv) intersect the position of the reads with annotated features from the corresponding
 687 reference sequence (**Table S2**). In order to calculate nucleotide percentage identity (%)
 688 between subgenomes for breakpoint windows and non-breakpoint windows, pairwise whole-
 689 genome alignments were obtained between *Saccharomyces* species corresponding to the
 690 species combination identified in our set of hybrids using Mugsy (v1.2.3)¹⁰⁷. Next, the 1kb
 691 interval coordinates from the alignments against the reference sequences were mapped on the
 692 whole-genome alignments and nucleotide percentage identity was calculated as number of
 693 matches on the total amount of bases in the window excluding gaps. We repeated the
 694 calculations for 500,100 and 50bp windows in order to identify microhomology regions
 695 within the starting 1kb windows. Then, we compared the identity distribution across
 696 breakpoint windows and non-breakpoint windows. The nucleotide percentage identity
 697 distribution across breakpoint windows and non-breakpoint windows was compared using the
 698 Wilcoxon Signed Rank Test implemented in the MASS package (v7.3-47)¹⁰⁸ in R. Twenty-
 699 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 inoculate the growth medium at an initial optical density at 600 nm (OD600) of 0.5 (roughly equivalent to 10⁷ 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, H₂S 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 H₂S 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 previously²⁸.

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 previously²⁸.

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 previously²⁸. 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 method¹¹⁰ 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*10⁷ 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*10⁶ cells mL⁻¹. The headspace of the samples was flushed with CO₂ 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. 2011¹¹. 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. 10⁸ 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

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

Biomass production in cold wort

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

Development and phenotypic evaluation of artificial hybrids

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

These hybrids were tested for biomass during wort fermentation at cold temperature as described in the previous section. The maltotriose that was still present after 7 days of fermentation was measured using the Dionex system (ICS 5000+, Thermo Fisher Scientific, USA). Fermentation samples were diluted in dH₂O, filtered (0.2 μm), and 10 μL were injected into the system. A Carboxypac PA20 column, kept at 30°C, was used, with a column flow of 0.3 $\mu\text{L min}^{-1}$. 250mM NaOH (Eluent 1) and 500 mM CH₃COONa + 100 mM NaOH (Eluent 2) were used as eluents.

822 **References**

- 823 1. Mallet, J., Besansky, N. & Hahn, M. W. How reticulated are species? *BioEssays* **38**,
824 140–149 (2016).
- 825 2. Abbott, R. *et al.* Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).
- 826 3. Shapiro, B. J., Leducq, J.-B. & Mallet, J. What Is Speciation? *PLOS Genet.* **12**,
827 e1005860 (2016).
- 828 4. Nolte, A. W. & Tautz, D. Understanding the onset of hybrid speciation. *Trends Genet.*
829 **26**, 54–8 (2010).
- 830 5. Baack, E. J. & Rieseberg, L. H. A genomic view of introgression and hybrid
831 speciation. *Curr. Opin. Genet. Dev.* **17**, 513–8 (2007).
- 832 6. Rieseberg, L. H., Archer, M. A. & Wayne, R. K. Transgressive segregation, adaptation
833 and speciation. *Heredity (Edinb)*. **83**, 363–372 (1999).
- 834 7. Lamichhaney, S. *et al.* Rapid hybrid speciation in Darwin’s finches. *Science* **359**, 224–
835 228 (2018).
- 836 8. Rieseberg, L. H. *et al.* Hybridization and the colonization of novel habitats by annual
837 sunflowers. *Genetica* **129**, 149–65 (2007).
- 838 9. Rieseberg, L. H. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence
839 from ribosomal genes. *Am. J. Bot.* **78**, 1218–1237 (1991).
- 840 10. Schwarzbach, A. E. & Rieseberg, L. H. Likely multiple origins of a diploid hybrid
841 sunflower species. *Mol. Ecol.* **11**, 1703–15 (2002).
- 842 11. Pryszcz, L. P. *et al.* The Genomic Aftermath of Hybridization in the Opportunistic
843 Pathogen *Candida metapsilosis*. *PLOS Genet.* **11**, e1005626 (2015).
- 844 12. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. Sequencing and
845 comparison of yeast species to identify genes and regulatory elements. *Nature* **423**,
846 241–254 (2003).
- 847 13. Greig, D. Reproductive isolation in *Saccharomyces*. *Heredity (Edinb)*. **102**, 39–44
848 (2009).

- 849 14. Liti, G., Barton, D. B. H. & Louis, E. J. Sequence diversity, reproductive isolation and
850 species concepts in *Saccharomyces*. *Genetics* **174**, 839–50 (2006).
- 851 15. Hou, J., Friedrich, A., De Montigny, J. & Schacherer, J. Chromosomal rearrangements
852 as a major mechanism in the onset of reproductive isolation in *Saccharomyces*
853 *cerevisiae*. *Curr. Biol.* (2014). doi:10.1016/j.cub.2014.03.063
- 854 16. Morales, L. & Dujon, B. Evolutionary role of interspecies hybridization and genetic
855 exchanges in yeasts. *Microbiol. Mol. Biol. Rev.* **76**, 721–39 (2012).
- 856 17. Naumov, G. I. *et al.* Natural polyploidization of some cultured yeast *Saccharomyces*
857 *sensu stricto*: Auto- and allotetraploidy. *Syst. Appl. Microbiol.* (2000).
- 858 18. Zhang, H., Skelton, A., Gardner, R. C. & Goddard, M. R. *Saccharomyces paradoxus*
859 and *Saccharomyces cerevisiae* reside on oak trees in New Zealand: evidence for
860 migration from Europe and interspecies hybrids. *FEMS Yeast Res.* **10**, 941–947
861 (2010).
- 862 19. Stefanini, I. *et al.* Social wasps are a *Saccharomyces* mating nest. *Proc. Natl. Acad.*
863 *Sci. U. S. A.* **113**, 2247–51 (2016).
- 864 20. Nakao, Y. *et al.* Genome sequence of the lager brewing yeast, an interspecies hybrid.
865 *DNA Res.* **16**, 115–29 (2009).
- 866 21. Walther, A., Hesselbart, A. & Wendland, J. Genome sequence of *Saccharomyces*
867 *carlsbergensis*, the world's first pure culture lager yeast. *G3 (Bethesda)*. **4**, 783–93
868 (2014).
- 869 22. Dunn, B. & Sherlock, G. Reconstruction of the genome origins and evolution of the
870 hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res.* **18**, 1610–23 (2008).
- 871 23. Bing, J., Han, P.-J., Liu, W.-Q., Wang, Q.-M. & Bai, F.-Y. Evidence for a Far East
872 Asian origin of lager beer yeast. *Curr. Biol.* **24**, R380-1 (2014).
- 873 24. Cavalieri, D., McGovern, P. E., Hartl, D. L., Mortimer, R. & Polsinelli, M. Evidence
874 for *S. cerevisiae* fermentation in ancient wine. *J. Mol. Evol.* **57 Suppl 1**, S226-32
875 (2003).
- 876 25. Michel, R. & McGovern, P. Chemical evidence for ancient beer. *Nature* 145–150

- 877 (1992).
- 878 26. McGovern, P. E. *et al.* Fermented beverages of pre- and proto-historic China. *Proc.*
879 *Natl. Acad. Sci.* **101**, 17593–17598 (2004).
- 880 27. Borneman, A. R. *et al.* Whole-genome comparison reveals novel genetic elements that
881 characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS*
882 *Genet.* **7**, e1001287 (2011).
- 883 28. Gallone, B. *et al.* Domestication and Divergence of *Saccharomyces cerevisiae* Beer
884 Yeasts. *Cell* **166**, 1397–1410.e16 (2016).
- 885 29. Gonçalves, M. *et al.* Distinct Domestication Trajectories in Top-Fermenting Beer
886 Yeasts and Wine Yeasts. *Curr. Biol.* **26**, 2750–2761 (2016).
- 887 30. Legras, J.-L. *et al.* Adaptation of *S. cerevisiae* to Fermented Food Environments
888 Reveals Remarkable Genome Plasticity and the Footprints of Domestication. *Mol.*
889 *Biol. Evol.* **35**, 1712–1727 (2018).
- 890 31. Peter, J. *et al.* Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates.
891 *Nature* **556**, 339–344 (2018).
- 892 32. Yue, J.-X. *et al.* Contrasting evolutionary genome dynamics between domesticated and
893 wild yeasts. *Nat. Genet.* (2017).
- 894 33. Dunn, B., Richter, C., Kvitek, D. J., Pugh, T. & Sherlock, G. Analysis of the
895 *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants
896 distributed in diverse yeast strains from differing industrial environments. *Genome*
897 *Res.* **22**, 908–24 (2012).
- 898 34. Duan, S. F. *et al.* The origin and adaptive evolution of domesticated populations of
899 yeast from Far East Asia. *Nat. Commun.* **9**, 2690 (2018).
- 900 35. Ludlow, C. L. *et al.* Independent Origins of Yeast Associated with Coffee and Cacao
901 Fermentation. *Curr. Biol.* **26**, 965–971 (2016).
- 902 36. Barbosa, R. *et al.* Multiple Rounds of Artificial Selection Promote Microbe Secondary
903 Domestication—The Case of Cachaça Yeasts. *Genome Biol. Evol.* **10**, 1939–1955
904 (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 Froberg 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. 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. Pérez-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

990 (2019).

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

993 72. Steensels, J. & Verstrepen, K. J. Taming Wild Yeast: Potential of Conventional and
994 Nonconventional Yeasts in Industrial Fermentations. *Annu. Rev. Microbiol.* **68**, 61–80
995 (2014).

996 73. Dzialo, M. C., Park, R., Steensels, J., Lievens, B. & Verstrepen, K. J. Physiology,
997 ecology and industrial applications of aroma formation in yeast. *FEMS Microbiol. Rev.*
998 **41**, S95–S128 (2017).

999 74. Mukai, N., Masaki, K., Fujii, T., Kawamukai, M. & Iefuji, H. *PADI* and *FDCI* are
1000 essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*.
1001 *J. Biosci. Bioeng.* **109**, 564–9 (2010).

1002 75. Mukai, N., Masaki, K., Fujii, T. & Iefuji, H. Single nucleotide polymorphisms of
1003 *PADI* and *FDCI* show a positive relationship with ferulic acid decarboxylation ability
1004 among industrial yeasts used in alcoholic beverage production. *J. Biosci. Bioeng.* **118**,
1005 50–5 (2014).

1006 76. Pasteur, L. Etudes sur la bière ses maladies, causes qui les provoquent, procédé pour la
1007 rendre inaltérable, avec une théorie nouvelle de la fermentation. 387 (1876).

1008 77. Teich, M. *Bier, Wissenschaft und Wirtschaft in Deutschland 1800-1914: ein Beitrag*
1009 *zur deutschen Industrialisierungsgeschichte.* (Böhlau, 2001).

1010 78. Hansen, E. Recherches sur la physiologie et la morphologie des ferments alcooliques
1011 V. Methodes pour obtenir des cultures pures de *Saccharomyces* et de mikroorganismes
1012 analogues. *C. R. Trav. Lab. Carlsb.* **2**, 92–105 (1883).

1013 79. Mertens, S. *et al.* A Large Set of Newly Created Interspecific *Saccharomyces* Hybrids
1014 Increases Aromatic Diversity in Lager Beers. *Appl. Environ. Microbiol.* **81**, 8202–
1015 8214 (2015).

1016 80. Peris, D. *et al.* Hybridization and adaptive evolution of diverse *Saccharomyces* species
1017 for cellulosic biofuel production. *Biotechnol. Biofuels* **10**, 78 (2017).

- 1018 81. Scannell, D. R. *et al.* The Awesome Power of Yeast Evolutionary Genetics: New
1019 Genome Sequences and Strain Resources for the *Saccharomyces sensu stricto* Genus.
1020 *G3 (Bethesda)*. **1**, 11–25 (2011).
- 1021 82. Liti, G. *et al.* High quality de novo sequencing and assembly of the *Saccharomyces*
1022 *arboricolus* genome. *BMC Genomics* **14**, 69 (2013).
- 1023 83. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler
1024 transform. *Bioinformatics* **25**, 1754–60 (2009).
- 1025 84. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing
1026 genomic features. *Bioinformatics* **26**, 841–842 (2010).
- 1027 85. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to
1028 single-cell sequencing. *J. Comput. Biol.* **19**, 455–77 (2012).
- 1029 86. Pryszcz, L. P. & Gabaldón, T. Redundans: an assembly pipeline for highly
1030 heterozygous genomes. *Nucleic Acids Res.* **44**, e113–e113 (2016).
- 1031 87. Kolmogorov, M., Raney, B., Paten, B. & Pham, S. Ragout-a reference-assisted
1032 assembly tool for bacterial genomes. *Bioinformatics* **30**, i302–i309 (2014).
- 1033 88. Wapinski, I., Pfeffer, A., Friedman, N. & Regev, A. Natural history and evolutionary
1034 principles of gene duplication in fungi. *Nature* **449**, 54–61 (2007).
- 1035 89. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local
1036 alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- 1037 90. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**,
1038 421 (2009).
- 1039 91. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
1040 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–80 (2013).
- 1041 92. Kück, P. & Meusemann, K. FASconCAT: Convenient handling of data matrices. *Mol.*
1042 *Phylogenet. Evol.* **56**, 1115–1118 (2010).
- 1043 93. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for
1044 automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*
1045 **25**, 1972–3 (2009).

- 1046 94. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
1047 large phylogenies. *Bioinformatics* **30**, 1312–3 (2014).
- 1048 95. Kozlov, A. M., Aberer, A. J. & Stamatakis, A. ExaML version 3: a tool for
1049 phylogenomic analyses on supercomputers. *Bioinformatics* **31**, 2577–9 (2015).
- 1050 96. Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E. &
1051 Stamatakis, A. How Many Bootstrap Replicates Are Necessary? *J. Comput. Biol.* **17**,
1052 337–354 (2010).
- 1053 97. R Development Core Team, R. R: A Language and Environment for Statistical
1054 Computing. *R Foundation for Statistical Computing* **1**, 409 (2011).
- 1055 98. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: an r package for
1056 visualization and annotation of phylogenetic trees with their covariates and other
1057 associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
- 1058 99. Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data integration using
1059 BEAST 1.10. *Virus Evol.* **4**, (2018).
- 1060 100. Raj, A., Stephens, M. & Pritchard, J. K. fastSTRUCTURE: variational inference of
1061 population structure in large SNP data sets. *Genetics* **197**, 573–89 (2014).
- 1062 101. Ayres, D. L. *et al.* BEAGLE: An Application Programming Interface and High-
1063 Performance Computing Library for Statistical Phylogenetics. *Syst. Biol.* **61**, 170–173
1064 (2012).
- 1065 102. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling
1066 trees. *BMC Evol. Biol.* **7**, 214 (2007).
- 1067 103. Drummond, A. J. & Suchard, M. A. Bayesian random local clocks, or one rate to rule
1068 them all. *BMC Biol.* **8**, 114 (2010).
- 1069 104. Worobey, M., Han, G.-Z. & Rambaut, A. A synchronized global sweep of the internal
1070 genes of modern avian influenza virus. *Nature* **508**, 254–7 (2014).
- 1071 105. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior
1072 Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **67**, 901–904
1073 (2018).

1074 106. Baele, G., Lemey, P. & Suchard, M. A. Genealogical Working Distributions for
1075 Bayesian Model Testing with Phylogenetic Uncertainty. *Syst. Biol.* **65**, 250–264
1076 (2016).

1077 107. Angiuoli, S. V. & Salzberg, S. L. Mugsy: fast multiple alignment of closely related
1078 whole genomes. *Bioinformatics* **27**, 334–342 (2011).

1079 108. Venables, W. N., Ripley B.D. Modern Applied Statistics With S. *Technometrics* **45**,
1080 111–111 (2003).

1081 109. Warnes, G. R. *et al.* gplots: Various R Programming Tools for Plotting Data. R
1082 package version 3.0.0. <http://CRAN.R-project.org/package=gplots>. (2016).

1083 110. Murtagh, F. & Legendre, P. Ward's Hierarchical Agglomerative Clustering Method:
1084 Which Algorithms Implement Ward's Criterion? *J. Classif.* **31**, 274–295 (2014).

1085 111. Calahan, D., Dunham, M., Desevo, C. & Koshland, D. E. Genetic analysis of
1086 desiccation tolerance in *Saccharomyces cerevisiae*. *Genetics* **189**, 507–519 (2011).

1087 112. Muir, A., Harrison, E. & Wheals, A. A multiplex set of species-specific primers for
1088 rapid identification of members of the genus *Saccharomyces*. *FEMS Yeast Res.* **11**,
1089 552–563 (2011).

1090 113. Pengelly, R. J. & Wheals, A. E. Rapid identification of *Saccharomyces eubayanus* and
1091 its hybrids. *FEMS Yeast Res.* **13**, 156–161 (2013).

1092

1093

Saccharomyces cerevisiae

Saccharomyces eubayanus

Saccharomyces uvarum

Saccharomyces kudriavzevii

Lineage

Interspecific hybrid type

a

b

c

d

a *S. cerevisiae* x *S. kudriavzevii*

b *S. uvarum* x *S. eubayanus*

c *S. cerevisiae* x *S. eubayanus*

a **b** **c** **d**

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