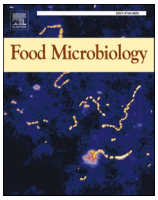
Food Microbiology 115 (2023) 104320

Contents lists available at ScienceDirect

Food Microbiology

journal homepage: www.elsevier.com/locate/fm



Do microbes evade domestication? - Evaluating potential ferality among diastatic *Saccharomyces cerevisiae*

Francisca Paraíso [a,b], Ana Pontes [a,b], Joana Neves [a,b], Kebaneilwe Lebani [c], Mathias Hutzler [d], Nerve Zhou [c], José Paulo Sampaio [a,b,*]

[a] UCIBIO – Applied Molecular Biosciences Unit, Department of Life Sciences, Nova School of Science and Technology, Portugal
[b] Associate Laboratory i4HB - Institute for Health and Bioeconomy, Nova School of Science and Technology, Portugal
[c] Department of Biological Sciences and Biotechnology, Botswana International University of Science and Technology, Private Bag 16, Palapye, Botswana
[d] Technical University of Munich, Research Center Weihenstephan for Brewing and Food Quality, Alte Akademie 3, 85354 Freising, Germany

ARTICLE INFO

Keywords:
Diastatic *Saccharomyces cerevisiae*
Saccharomyces diastaticus
Brewery microbial contaminants
Beer microbiology
STA1 gene
Feral

ABSTRACT

Certain lineages of the wine, beer and bread yeast *Saccharomyces cerevisiae* have diastatic activity. They contain the chimeric gene *STA1* that codes for an extracellular glucoamylase which enables the strains to degrade starch and dextrins. Beer contaminations by diastatic yeasts can be dangerous because they can cause super-attenuation due to the consumption of otherwise non-fermentable oligosaccharides, gushing and off-flavours. Given that diastatic yeasts can be used for beer fermentation it is important to understand the relationship between production and contaminant strains, their natural reservoirs and entry routes into the brewery. Here, we analyze real cases of contamination in a Portuguese craft brewery over a period of 18 months. By analyzing with whole genome sequencing several contaminants, we show that recurrent contaminations by diastatic yeasts are caused by environmental strains. Moreover, some beer contaminants were closely related to diastatic environmental strains isolated in Botswana. We observed the widespread presence of domestication signatures in diastatic strains. Moreover, the combined phylogeny of *STA1* and its ancestor, *SGA1*, suggested a single *STA1* origin, as ancient as the entire lineage of diastatic yeasts. Together, our results suggest that diastatic yeasts isolated in natural settings could be escaping from domestication settings and becoming feral.

1. Introduction

Two species of the genus *Saccharomyces* are of great importance for the production of major beer types. *S. pastorianus* is important for the fermentation of lagers, while *S. cerevisiae* is important for the fermentation of ales, stouts and wheat beers, among other types. Recent population, evolutionary and comparative genomics studies have revealed that *S. cerevisiae* is composed of multiple populations and that some of them are clearly associated with different types of fermentations (e.g. Gallone et al., 2016; Duan et al., 2018; Peter et al., 2018; Bigey et al., 2020). As such, it is now possible to recognize the Wine, Bread and Sake populations, among many others (Pontes et al., 2020). For beer, two main *S. cerevisiae* populations have been identified, the Beer 1 and Beer 2 populations (Gallone et al., 2016). A third group designated African Beer has been recently analysed and compared with the other two populations (Saada et al., 2022). Contrary to Beer 1 and Beer 2 strains, African Beer strains are not implicated in the types of industrial fermentations that yield a beverage typically recognized as "beer". Instead, they participate in the fermentation of different artisanal African beverages based on cereals such as sorghum, millet, and cassava as well as non-starchy raw materials such as wild fruits and plant exudates. Most types of English-Irish ales and German-style alt, kölsch and wheat beers, together with some Belgian beers, are mainly fermented by strains of the Beer 1 population, whereas saison-type beers and also different types of ales, are fermented by a distinct population designated Beer 2 (Gallone et al., 2016; Gonçalves et al., 2016). These two populations are distantly related, and the Beer 2 lineage is more closely related to the Wine lineage than to the Beer 1 lineage (Gallone et al., 2016; Pontes et al., 2020). This indicates that the genetic and phenotypic changes associated with domestication of beer yeasts occurred independently at least two times.

The utilization of maltotriose, one of the most abundant sugars in beer wort, illustrates the separate origins of Beer 1 and Beer 2 yeasts. In the Beer 1 population, an allele of the *MAL11* gene (*AGT1*), encodes a

* Corresponding author. Department of Life Sciences, Faculdade de Ciências Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
E-mail address: jss@fct.unl.pt (J.P. Sampaio).

https://doi.org/10.1016/j.fm.2023.104320
Received 23 February 2023; Received in revised form 1 June 2023; Accepted 7 June 2023
Available online 7 June 2023

high affinity maltotriose transporter, normally not found in other *S. cerevisiae* lineages, but present in this lineage and in multiple copies (Gonçalves et al., 2016). The strains of the Beer 2 population are also able to ferment maltotriose but the *AGT1* allele is either absent or non-functional. Beer 2 population strains contain the *STA1* gene, which is associated with the extracellular hydrolysis of maltotriose. The *STA1* gene is absent in the Beer 1 lineage. (Krogerus et al., 2019). *STA1* codes for an extracellular glucoamylase and appears to be chimeric, consisting of rearranged gene fragments from *FLO11*, linked to flocculation, and *SGA1*, that encodes a intracellular sporulation-specific glucoamylase (Yamashita et al., 1987). Moreover, detailed analyses have shown that some strains from the Beer 2 population lack the *STA1* gene (Krogerus et al., 2019; Pontes et al., 2020) and that *STA1*-positive strains have a variable diastatic activity. Variation in diastatic activity is also caused by differential gene expression given that poorly diastatic strains have a deletion in the promoter region of *STA1*. Therefore, the Beer 2 population encompasses a widely diverse group of strains with respect to diastatic activity, ranging from highly diastatic strains to some with very low or null activity (Meier-Dörnberg et al., 2018).

The role of *STA1* in the extracellular degradation of starch and oligosaccharides in the brewery is known for more than five decades (Andrews and Gilliland, 1952) and is associated with the currently obsolete names *S. diastaticus* or *S. cerevisiae* var. *diastaticus* (Pontes et al., 2020). When such a *S. cerevisiae* strain gains access to beer, it causes super-attenuation by consuming otherwise non-fermentable oligosaccharides, thus yielding increased CO_2 and ethanol levels, a drier mouthfeel and even off-flavours (Andrews and Gilliland 1952; Hutzler et al., 2012; Meier-Dörnberg et al., 2018). If the diastatic strain develops in packaged beer, "gushing" is a common outcome. As such diastatic *S. cerevisiae* strains are seen by brewers as dangerous contaminants (Hutzler et al., 2012; Suiker and Wösten, 2022).

The link between diastatic strains and Beer 2 producing strains was established only recently following genomic analyses (Krogerus et al., 2019). Therefore, to provide a uniform designation for the ever-increasing number of *S. cerevisiae* populations, we proposed the "Beer 2 – Diastaticus" designation for this group (Pontes et al., 2020), that here we shorten to "Beer 2D". It is now evident that this distinct population of *S. cerevisiae* encompasses at least two types of strains – those that can be used for brewing and those that can cause beer deterioration (Meier-Dörnberg et al., 2018). This better understanding of diastatic yeasts still faces important open questions. First, given that contamination by diastatic yeasts is a problem that affects commercial breweries (Meier-Dörnberg et al., 2017), especially the smaller ones that normally have less stringent sanitation procedures, it is relevant to understand the sources and routes of contamination. Given that diastatic yeasts can be used for beer fermentation, it can be hypothesized that contaminations in the brewery are caused by commercial strains. However, it can also be hypothesized that the contaminants are environmental *S. cerevisiae* strains with diastatic capacity. Indeed, the recent finding of *STA1*, the gene associated with diastatic activity, in a *S. cerevisiae* population found in French Guiana (Krogerus et al., 2019) could be an indication of the occurrence of diastatic yeasts in natural environments. These yeasts belong to a population that is distinct from the Beer 2D population and were isolated from *cachiri*, a starch-rich fermented beverage consumed by Wayampi Amerindians as well as from their stool samples after the consumption of the beverage (Angebault et al., 2013).

Here, we take advantage of the new understanding of diastatic yeasts at the genome level to study real cases of contamination in a Portuguese craft brewery. Recurrent contaminations with suspected diastatic yeasts were detected, representative cultures were isolated and used for whole-genome sequencing. We found that the contaminants did not derive from the diastatic commercial culture used previously in the brewery and that three genetically distinct contaminants colonized and persisted in the brewery. By analyzing diastatic yeasts from different provenances and by studying the phylogeny of *STA1* together with the phylogeny of its ancestor, the gene *SGA1*, we evaluate the possibility that diastatic yeasts are escaping from the domestication settings where they arose and becoming feral, thus explaining why the genomes of some contaminants are similar to the genomes of diastatic yeasts isolated from natural settings.

2. Methods

2.1. Isolation of diastatic S. cerevisiae

Isolations carried out from fermented beverages and wild fruits in Botswana were conducted at 25 °C by direct inoculations of samples from fermentations on YMA (1% glucose, 0.5% peptone, 0.3% yeast extract, 0.3% malt extract and 2% agar) or by enrichment in YM supplemented with 500 ppm chloramphenicol followed by isolation on YMA for fruit samples.

Beer analyses were carried between the first and third week after bottling or kegging (beer batches that yielded negative results for diastatic yeasts were re-examined after 6–10 weeks). The analyses included measurements of ºBrix and pH, recording of gushing, and a microscopic examination. Beer samples were directly inoculated (or diluted when necessary) on YMA (1% glucose, 0.5% peptone, 0.3% yeast extract, 0.3% malt extract and 2% agar). When low levels of cells were observed in the microscope, an enrichment step was performed. In those cases, 5 mL of beer were inoculated in 50 mL of YPD (1% yeast extract, 2% peptone, 2% glucose) and incubated overnight at 25 °C with shaking (150 r.p.m.) prior to inoculation on YMA. Cell counts in beer samples were determined by counting the colonies that grew on YMA. For the search of the diastatic genotype by PCR, 5% of the colonies of each positive sample (approximately 5–30) were transferred to new YMA plates for purification.

2.2. Multiplex PCR for the detection of STA1 and the deletion of STA1 promoter and Sanger sequencing of STA1

The presence of *STA1* gene and of the intact *STA1* promoter were tested with a colony picking multiplex PCR protocol. We used primers STA_RT_2_FW and STA_RT_2_RV, designed by us to amplify the *STA1* gene, and primers STA1_UAS_Fw and STA1_UAS_Rv from (Krogerus et al., 2019) to amplify the intact *STA1* promoter region (Table S1). PCR reactions were carried out with DreamTaq Polymerase (Thermo Scientific) and primer concentrations were 0.2 μM. The following PCR program was used: 95 °C for 5min, (95 °C for 20 s, 55 °C for 30 s, 72 °C for 20 s) × 30 cycles, 72 °C for 2 min. PCR products were separated by electrophoresis in a 2% agarose gels. To control the experiment, we amplified the actin gene *ACT1* with primers, ACT1_Fw and ACT1_Rv (Table S1) with the following PCR program: 95 °C 5min, (95 °C 15 s, 55 °C 30 s, 72 °C 45 s) × 35 cycles, 72 °C 4 min. PCR products were also separated as above.

For Sanger sequencing of *STA1,* the gene was initially amplified by PCR using primers STA1_Complete_Fw and STA1_Complet_Rv (Table S1), producing a 2468 bp amplicon (DreamTaq Polymerase and primer concentrations of 0.2 μM). The following PCR program was employed: 95 °C 5min, (95 °C 45 s, 52 °C 45 s, 72 °C 2min) × 30 cycles, 72 °C 5 min. PCR products were purified using Illustra GFX PCR DNA and Gel Band Purification kit. Sanger sequencing was performed at STAB VIDA (Caparica, Portugal), with primers STA1_Complete_Fw and STA1_Complet_Rv. Due to the PCR product size and in order to fully sequence this gene, 2 μL of the previously obtained amplicon (10 μg/μL) were used to amplify the central region of *STA1*. PCR reactions were carried out with primers STA1_Center_Fw and STA1_Center_Rv (Table S1) (DreamTaq Polymerase and primer concentrations of 0.2 μM), producing an 662 bp amplicon with the following program: 95 °C 5min, (95 °C 30 s, 59 °C 30 s, 72 °C 30 s) × 30 cycles, 72 °C 4 min. PCR products were purified and sequenced as above using primers STA1_Center_Fw and STA1_Center_Rv. All sequences were aligned with muscle

in MEGA7 (Kumar et al., 2016).

2.3. Genome sequencing, read alignment and genotype calling

For genome sequencing, DNA was extracted from overnight grown cultures and paired-end Illumina NextSeq (300 cycles) reads were obtained. Chromosomal nucleotide polymorphisms (SNPs) from the whole genome were extracted following an adapted GATK germline short variant discovery pipeline (Poplin et al., 2017). Sequenced reads for each isolate were mapped to a combined reference genome of *S. cerevisiae* S288C (version R64-1-1) and *S. paradoxus* CBS 432 (ASM207905v1) using BWA v.0.7.17 (Li and Durbin, 2009) and duplicated reads were marked with Picard v.2.22.8 (http://broadinstitute.github.io/picard/). SNP, INDEL and genotyping determination was performed on all samples simultaneously using a local re-assembly of haplotypes (GATK HaplotypeCaller, Genomics DBImport, and Genotype GVCFs) and standard hard filtering parameters of variant quality scores recalibration were adjusted according to GATK best practices recommendations (GATK Variant Filtration with parameter values QD < 2.0, QUAL <30.0, SOR >3.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0) (Depristo et al., 2011). The final SNPs matrix included 270577 high-quality homozygous SNPs across 100 strains included in this study.

2.4. Phylogenetic analyses, survey of specific genes and divergence analyses

The main phylogeny was constructed using the maximum-likelihood method as implemented in IQ-TREE v. 16.11 (Nguyen et al., 2015), using the best estimated model of sequence evolution, TVM + F + ASC + G4, and the ultrafast bootstrap approximation with 1000 replicates (Minh et al., 2013). The software iTOL v5 (Letunic and Bork, 2021) was used for visualization. Single gene phylogenies were inferred in MEGA7 (Kumar et al., 2016) using the maximum-likelihood method and Kimura 2 parameter model.

For the investigation of genes of interest, short reads were processed with trimmomatic v. 0.36 (Bolger et al., 2014) to remove adapter sequences and for whole genome assemblies SPAdes v. 3.13.1 (Bankevich et al., 2012) was used. For each genome a local database was prepared and genes of interest were retrieved by searching ORFS with BLASTn. The presence of *STA1* and its promoter region was investigated using queries from Krogerus et al. (2019). In addition, long read sequences taken from public databases were assembled using Canu v2.2 (Koren et al., 2017) and used to retrieve the sequence of *STA1* by employing the GenBank sequence X02649.1 as query. *SGA1, PRK1, FYV10, FCM1* and *MRS1* were investigated using sequences from S288C as queries (retrieved from SGD database). Furthermore, *AQY1* and *AQY2* were investigated using sequence from YPS163 (Will et al., 2010). Finally, for the investigation of regions A, B and C, queries with all genes of each region from EC1118 (Novo et al., 2009) were produced, and the presence/absence of each individual gene was assessed. The percentage of heterozygosity was assessed based on Peter et al. (2018) and Duan et al. (2018). Gene sequence divergence was estimated using Variscan v2.0 (Hutter et al., 2006), with parameters CompleteDeletion = 0, FixNum = 1 and NumNuc = 4. To analyze the similarity degree of closely related genomes we used MUMmer utility *dnadiff* with default settings (https://github.com/marbl/MUMmer3/blob/master/docs/dnadiff.README), that uses pairs of assembled genomes as input to calculate SNPs and INDELS.

3. Results

3.1. Evidence for multiple and independent occurrences of contaminant diastatic yeasts

Between May 2020 and October 2021, we analysed 41 batches of bottled beers produced in a Portuguese craft brewery in order to investigate the occurrence of diastatic contaminants (Table S3). Prior to the beginning of the analysis, in February of 2020, a beer fermentation had been prepared with a commercial culture of a diastatic yeast. An additional utilization of this type of commercial yeast occurred during the period of our study, in July of 2020. In June 2021 the brewery plant was moved to another location. From 41 samples that were analysed we isolated *Saccharomyces*-like cultures in 37 of them. However, when surveyed for the presence of the *STA1* gene, the diagnostic feature of a *S. cerevisiae* diastatic strain, 24 isolates were positive and 13 were negative (Table S3). The negative results could be attributed to the detection, in the finished beer, of the original commercial culture, or of a non-diastatic *S. cerevisiae* contaminant. The remaining *STA1* positive cultures were regarded as contaminant diastatic *S. cerevisiae*.

Next, we selected five diastatic cultures, representing the temporal span of the survey, for whole-genome sequencing. We also sequenced the production diastatic culture. The phylogenomic placement of the five contaminants is shown in Fig. 1. In the phylogeny, we used a representative number of reference genome sequences from the Beer 2D population that were obtained from production brewing strains, including the diastatic production strain used in the brewery, brewing contaminants and putative wild strains (Table S2). Moreover, we included representatives of additional *S. cerevisiae* populations known to be related to the Beer 2D population (Peter et al., 2018; Pontes et al., 2020). Unsurprisingly, the five contaminants were placed in the Beer 2D population. They formed two separate clusters that did not include the production diastatic strain used in the brewery. This clearly showed that the production strain was not implicated in the contaminations detected in the brewery.

Since the designations of the contaminants follow a chronological order, it is possible to conclude that contaminant 1.3 (May 2020) and contaminant 4.1 (August 2020) represent clearly distinct and therefore independent contamination events because they are placed apart in the phylogenetic tree. On the contrary, contaminants 4.1 (August 2020) and 10.1 (early November 2020) appear to be undistinguishable in the phylogeny. However, the single nucleotide polymorphisms (SNPs) matrix upon which the phylogeny of Fig. 1 was generated contained 12 SNPs that separated these two contaminants (Supplementary Table S4). Moreover, the comparison of the genome guided assemblies of each of these two strains against the reference *S. cerevisiae* genome yielded more than 10.000 SNPs separating them, whereas equivalent assemblies of a control production strain sequenced twice by us had 1664 SNPs (Supplementary Table S4). This suggests that the two strains, albeit being genetically very similar, are not identical. Nevertheless, for 4.1. and 10.1 we believe it is more prudent to refrain from a definitive assessment on a single or double contamination, noting instead their genetic closeness. The same situation was observed for the contaminants detected in May 2020 (1.3) and in late November 2020 (11.3) that differed by 8 SNPs in the matrix used to construct the phylogeny of Fig. 1. The last contaminant studied (20.1, October 2021) clustered with contaminants 1.3 and 11.3 but is more distantly related (Fig. 1). Since this last contaminant was found after the brewing plant was moved to a new location and new pipes and hoses were installed, we think it should be regarded as a distinct contaminant. It thus appears that the contaminants that we detected in the brewery originate from two clearly distinct genetic stocks, one corresponding to 4.1 and 10.1 and the other to 1.3, 11.3 and 20.1.

Contaminants 4.1. and 10.1 were also very similar to two other strains isolated by us in Botswana. One of these strains (N14) was isolated from wild fruits of the marula tree (*Sclerocarya birrea* subsp *caffra*). This tree is locally known as elephant tree because elephants are attracted to its fruits, supposedly becoming inebriated after eating the ripened and fermenting fruits (Morris et al., 2006). The other strain (Ma01) was isolated from *khadi*, a traditional beverage resulting from the spontaneous fermentation of wild fruits of *Grewia flava,* a common shrub in Southern Africa (Motlhanka et al., 2020). The finding of

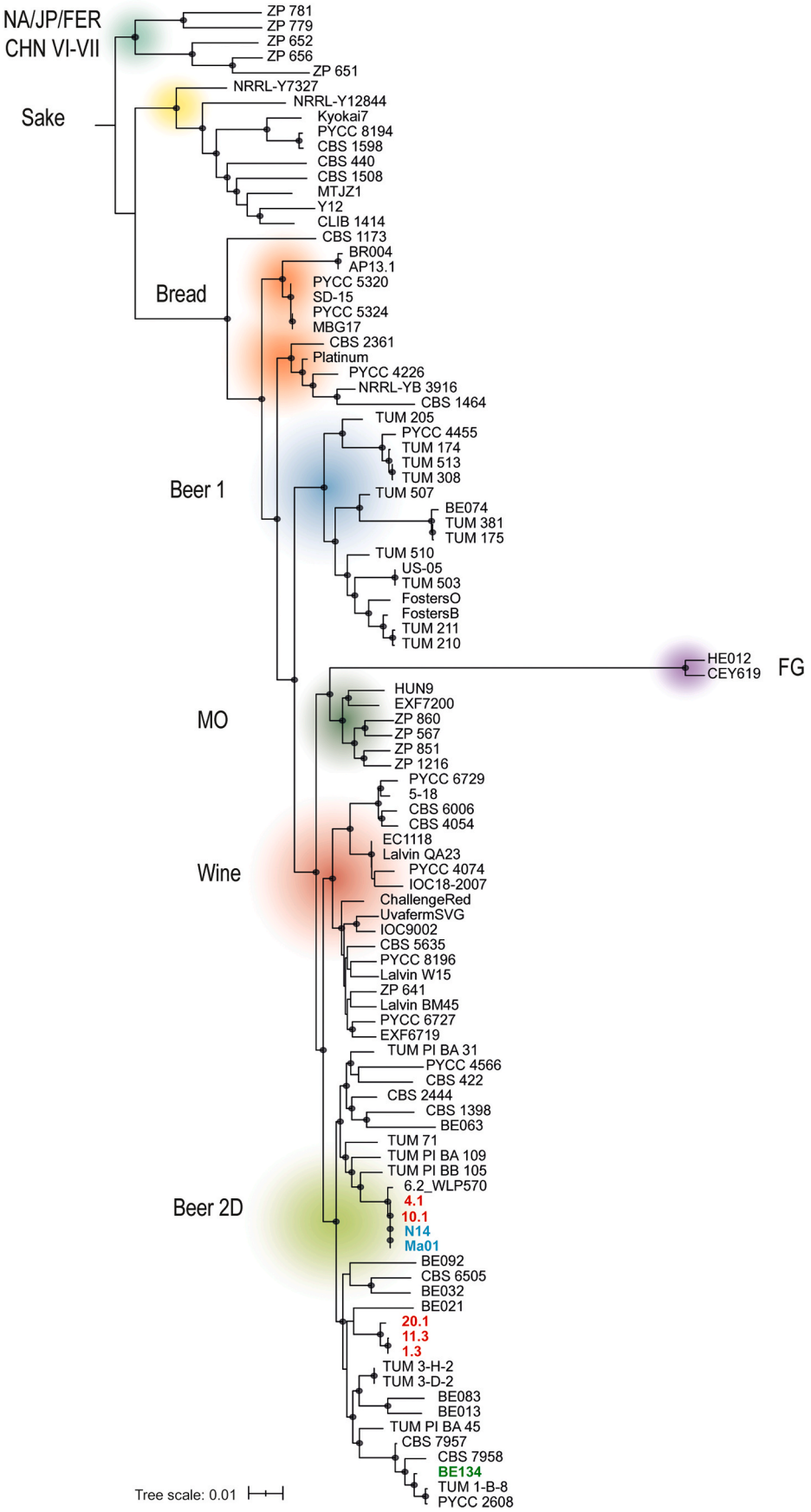


Fig. 1. Diastatic (*STA1* positive) contaminants of craft beer belong to Beer 2D population and are distinct from production strains. Phylogeny constructed from 100 genome sequences and 270577 high-quality homozygous single nucleotide polymorphisms, by applying the Maximum Likelihood method and the TVM + F + ASC + G4 model of sequence evolution as inferred in IQ-TREE. The brewery contaminants are indicated in red, the diastatic production strain used in the craft brewery is indicated in green, and the strains isolated in Botswana from *Marula* fruit and *khadi* that are similar to the brewing contaminants are indicated in blue. The designations of the various populations are indicated (MO, Mediterranean Oaks; NA/JP/FER/CHN VI-VII, North America – Japan - Far East Russia - China VI-VI). The phylogeny was rooted with members of the NA/JP/FER/CHN VI-VII population. Black dots in tree nodes depict bootstrap support values above 95% (1000 replicates) and branch lengths correspond to the predicted number of substitutions per site. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

brewery contaminants in Portugal with genomes very similar to putative wild strains in Botswana was unexpected. Among these four genomes no identical pair was found. The lowest number of SNPs was two (SNPs matrix) and was observed in the comparison of 10.1 with Ma01 (Table S4). These two strains differed by 10182 SNPs when their guided assemblies were compared.

3.2. *Contaminant diastatic yeasts are transitioning from the domesticated state to the feral state*

The remarkable similarity between some of the brewing contaminants and strains isolated in Boswana contrasted with the finding that diastatic contaminant yeasts were unrelated to the diastatic production strain used in the brewery. These observations prompted us to explore two alternative hypotheses in order to investigate in more detail the possible causes of the association between contaminants that colonize the brewery environment and strains found in more natural settings. The "wild origin" hypothesis posits that these contaminants are wild yeasts, in the sense that they colonize the brewery from outside natural sources; these environments thus represent the natural reservoir and therefore the niche of wild diastatic yeasts. The "feral" hypothesis, on the contrary, proposes that the contaminants derive from domesticated production strains that escaped fermentation, similarly to feral animals or plants that are found in the wild but descend from domesticated ancestors. Although both hypotheses predict that diastatic yeasts have a global distribution not circumscribed to the brewing environment, the "wild origin" hypothesis predicts that isolates obtained in natural environments do not have domestication signatures whereas the "feral" hypothesis predicts the opposite.

Finding genomes similar to the brewing contaminants in wild fruits from pristine and protected game reserves and spontaneous fermentations in Botswana appeared to support the "wild origin" hypothesis. However, to ascertain that these contaminants are truly wild it is necessary to confirm that signatures of domestication are absent in these strains. In previous publications we have documented (or expanded on documented cases identified by others) several domestication signatures. One case concerns the presence of regions A, B, and C, three genomic regions acquired from non-*Saccharomyces* yeasts and relatively frequent in wine strains (Novo et al., 2009). Another case relates to the inactivation of *AQY1* and *AQY2*, two paralogous genes that code for aquaporins, water channels that, when functional, tend to be detrimental in domesticated wine or beer strains given the high osmolarity of wine and beer must (Gonçalves et al., 2016; Will et al., 2010). A third case concerns the accumulation of heterozygous sites along the genome, which is typically increased in domesticated strains (Peter et al., 2018).

Using the tree topology of Beer 2D yeasts of Fig. 1, we added

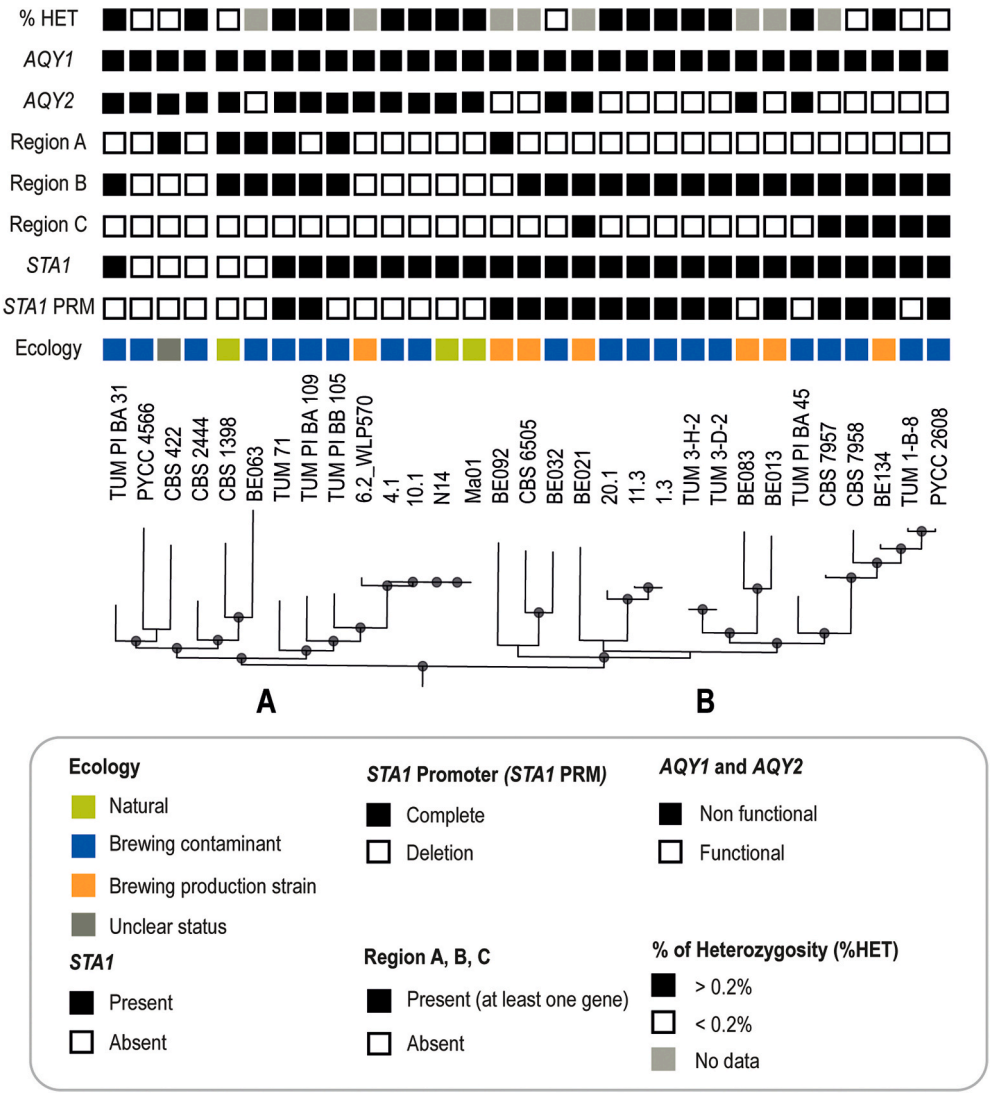


Fig. 2. The Beer 2D population does not contain truly wild strains. Relevant genetic and ecological characteristics are shown for the strains of this population using the topology of the tree of Fig. 1.

information on the presence or absence of *STA1* for each strain and ascertained whether its promoter region was complete or not. We also determined the presence of regions A, B and C; the inactivation of *AQY1* and *AQY2*; and the degree of heterozygosity (Fig. 2 and Table S6). Finally, we classified each strain with respect to its isolation source, i.e. distinguished strains that were isolated from a natural environment from those obtained from the brewing environment, differentiating in this case production and contaminant strains (Fig. 2.). First, we noted that Beer 2D yeasts are divided into two subclades that in general match what was found in an earlier study involving less genomes (Pontes et al., 2020). Given the widespread distribution of regions A, B and C, (most of the times the complete region was retrieved but, in some cases, a lower number of genes from each of these regions was found, as shown in Table S6), inactivation of *AQY1* and *AQY2,* and cases of increased heterozygosity, it is reasonable to assume that the entire clade represents a domesticated lineage. Moreover, beer contaminant strains are equally distributed in both subclades, some having the *STA1* gene and its promoter region complete, while others lack one or both elements. In subclade A, absence of *STA1,* or of its promoter region in cases where *STA1* is present, was observed in 12 of the 14 strains in this clade (86%). The opposite situation occurred in subclade B where that condition occurs in only three of the 17 strains (18%). The three strains associated with natural environments were only found in subclade A, whereas beer production strains were predominantly found in subclade B (six out of seven). It thus appears that subclade B is enriched in strains with the complete genetic makeup necessary for the typical diastatic phenotype whereas in subclade A the potential for diastatic activity is more limited. As already mentioned, the five contaminant strains isolated by us

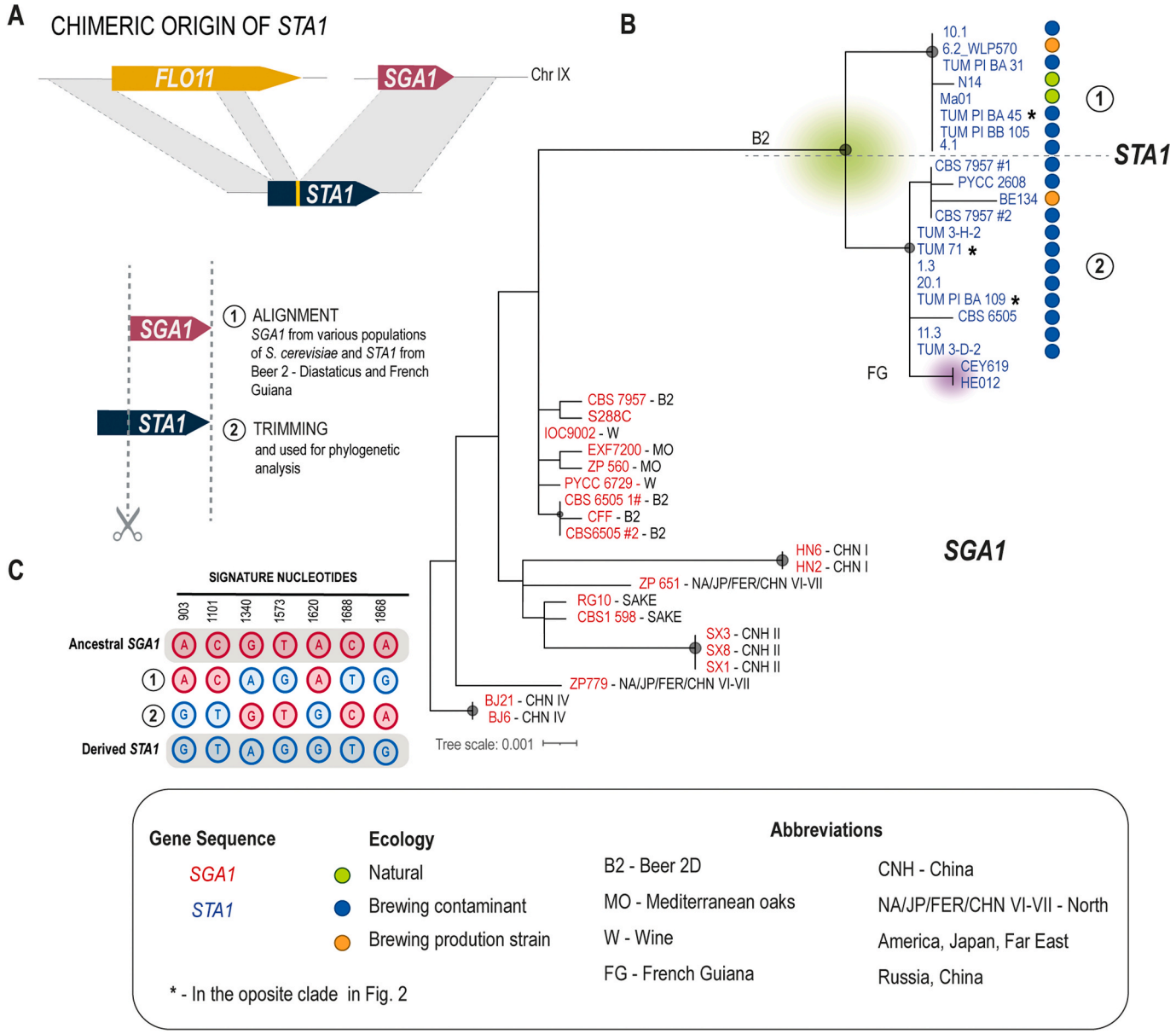


Fig. 3. Combined phylogenetic analysis of *SGA1* (complete gene) and *STA1* (catalytic domain only). (**A**) Representation of the chimeric origin of *STA1* (adapted from Krogerus and Gibson 2020) and of our strategy for sequence treatment for phylogenetic analysis. (**B**) Phylogeny combining *SGA1* sequences from wild and domesticated populations (S288C, reference retrieved form SGD) and *STA1* from representatives of the Beer 2D and French Guiana populations (Supplementary Data 1). The phylogeny was constructed using the Maximum Likelihood method and the Kimura −2 parameter model, was based on 53 informative sites and was rooted with the *SGA1* sequence of *S. paradoxus* CBS 432, not shown (accession XM_033911064.1, retrieved form NCBI). (**C**) Signature nucleotides of the ancestral *SGA1* sequence (nucleotides shown in red) and the derived *STA1* gene (nucleotides shown in blue). The ancestral or derived substitutions fixed in subclades 1 and 2 are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

formed two separate groups, one belonging to subclade A and another one belonging to subclade B. The strains in subclade A shared a close resemblance with Botswanan strains which included the absence of the promoter region of *STA1,* absence of regions A, B and C and inactivation of aquaporin genes *AQY1* and *AQY2* by the same mutations (Fig. 2 and Table S6). The other contaminant strains (subclade B) had the *STA1* promoter region, contained region B, and had a functional *AQY2* gene.

The presence of multiple domestication signatures in all stains is the expected outcome of the "feral" hypothesis, but not of the "wild origin" hypothesis. It is important to stress that wild populations like those found in Asia (Lee et al., 2022; Wang et al., 2012) or Southern Europe (Almeida et al., 2015) do not possess the domestication signatures investigated here (Pontes et al., 2020). Therefore, we suggest that the contaminants are formerly domesticated brewing strains that, given their widespread occurrence, even in substrates not related to beer brewing, are becoming feral.

3.3. Origin and divergence of STA1

Given that *STA1*, the gene coding for diastatic activity, is the product of a gene fusion event between *FLO11* and *SGA1* (reviewed in Krogerus and Gibson 2020), we sought to compare, in a phylogeny, not only *STA1* sequences from different strains harbouring this gene but also *SGA1* sequences from representatives of different populations of *S. cerevisiae*. We reasoned that such a comparison could provide clues on the origin and diversity of *STA1*. Having a more precise understanding of these questions could in turn contribute to a better understanding of the emergence of ferality. Our sequence analyses focused solely on the 3′ end of *STA1*, which is homologous to *SGA1* (Fig. 3A). This region contains the catalytic domain of *STA1*, and codes for a glucoamylase. Whereas *SGA1* codes for an intracellular enzyme, the fusion that gave rise to *STA1* leads to the production of an extracellular enzyme. This happens because the 5′ end of *STA1* is homologous to *FLO11*, a gene implicated in flocculation. The *FLO11*-derived peptide allows for extracellular secretion of the glucoamylase derived from *SGA1* (Adam et al., 2004).

For this analysis we used 19 *SGA1* sequences directly retrieved from short-read genome data (Table S2). The same approach could not be used for *STA1* given its chimeric nature and difficulty in distinguishing between *STA1, SGA1* and *FLO11* based on assemblies generated from short reads. Therefore we used three *STA1* sequences retrieved from long read genome data (Saada et al., 2022) and Sanger sequenced 19 *STA1* genes (Table S2), as explained in the methods section. A phylogeny of *SGA1* and the catalytic domain of *STA1* is shown in Fig. 3B. One branch of the phylogeny contains only *SGA1* sequences that were found in representatives of different populations of *S. cerevisiae* such as Asian wild and domesticated populations (China and Sake, respectively), the Wine population and their closest wild relatives (Mediterranean Oaks), and *SGA1* sequences from members of the Beer2D population. The other branch of the phylogeny contains *STA1* sequences, which are subdivided in two subclades, 1 and 2, that for the most part correspond to subclades A and B of Fig. 2, respectively. The similarity between the evolutionary relationships within the Beer2D population (Fig. 2) and of *STA1* (Fig. 3B) suggests that *STA1* was already present when this lineage was formed. Although the resolution in the phylogeny is poor, the *SGA1* sequences that appear to be closer to *STA1* are those of the Wine and Beer 2D population. The *STA1* sequences representing the French Guiana population are unique since they constitute the only case known to date of occurrence of *STA1* outside the Beer2D population. Despite being found in a distinct population so far only found in Central America, these sequences belong clearly to subclade 2 of *STA1* in the SGA1 – STA1 phylogeny, albeit in a slightly isolated position (Fig. 3B). The *STA1* sequences have a nucleotide diversity (π) of 0.00309 which is comparable with the diversity of the Beer2D population (0.00295, according to Gallone et al., 2016). These similar values support the scenario of the origin of *STA1* being associated with the emergence of the population. For *SGA1* we measured a global, species level, diversity of 0.00693 which is comparable to the diversity measured for the species *S. cerevisiae* (0.00663, according to Duan et al., 2018) (Table S7). We also surveyed the 9-bp stretch of similar *SGA1- FLO11* sequences that is implicated in the fusion of the two genes and found no variation among the studied *STA1* sequences (Fig. S1), which is also an indication of a possible single origin. We analysed in detail all sequences used in the phylogeny of Fig. 3B regarding signature substitutions and found seven substitutions along the sequence that are specific to subclades 1 or 2. Together, they separate *STA1* from the ancestral *SGA1*-like sequence (Table S8). We also investigated four flanking genes (*FMC1, FYV10, MRS1* and *PRK1*) of *SGA1, FLO11* and *STA1* in deeply sequenced genomes that contain both *SGA1* and *STA1* (Saada et al., 2022). The individual phylogenies of those genes clearly separated the *SGA1* and *FLO11* neighbors from the *STA1* neighbors (Fig. S2), thus supporting a non-recent origin of *STA1* that allowed for the observed divergence.

Taken together these results indicate that: (i) the event that led to the emergence of *STA1* via duplication, non-reciprocal translocation and subsequent gene fusion is sufficiently old to allow for the observed divergence of *STA1* and its flanking genes from their ancestor genes; (ii) the *STA1* allele found in the French Guiana population was likely acquired from the Beer 2D population; (iii) although definitive evidence is lacking, it is possible that *STA1* was formed in a single event linked to the origin of the Beer2D lineage. Therefore, subclades 1 and 2 seen in the *STA1* phylogeny could be an indication of an early divergence after a single gene fusion event or of two independent events that gave rise to the existing versions of *STA1*.

4. Discussion

Diastatic *Saccharomyces* strains are important beer spoilage agents, being able to metabolize dextrins and starch, thus producing super-attenuation, weakened body, increased alcohol content, over carbonation and even off-flavours (Andrews and Gilliland 1952; Hutzler et al., 2012; Meier-Dörnberg et al., 2018). Moreover, although recent studies have suggested a considerable incidence of spoilage with diastatic yeasts (Meier-Dörnberg et al., 2017; Latorre et al., 2020), little is known on the natural reservoirs and entrance routes of diastatic yeasts into the brewery. Here we show that beer contaminations by diastatic yeasts are caused by environmental strains of *S. cerevisiae* and not by the production strain used for brewing. We also show that recurrent contaminations by diastatic yeasts over a period of 18 months in the same brewery were caused by two clearly distinct genetic stocks which could represent up to five strains.

We suggest that the relatively frequent contamination events observed in our study might normally remain unnoticed by craft brewers for two reasons. First, we employed an isolation and diagnosis protocol specific for the detection of diastatic yeasts that involved a multiplex PCR for the detection of *STA1* and the deletion of its promoter region. This allowed a very accurate identification of diastatic yeasts that is not common practice among craft brewers. We were therefore able to detect and isolate diastatic yeasts from bottled beer batches for which off flavours or gushing had not been reported. Secondly, since our study was undertaken during one of the most severe periods of the covid-19 pandemic in Europe, beer consumption was reduced, and batches of bottled beers remained in the brewery warehouse for abnormally long periods. This naturally favoured the development of diastatic contaminants that are well adapted to the conditions of finished bottled beer.

By integrating the evidence on the occurrence of diastatic contaminants with whole-genome data from multiple lineages of *S. cerevisiae* we were able to identify the closest relatives of the brewery contaminants. Surprisingly, they belonged to a group of *S. cerevisiae* strains isolated in Botswana from fruits of the marula tree and from *khadi*, a wild fruit-based spontaneously fermented beverage. Thus, we show that rather than being restricted to the brewing environment, diastatic yeasts are widespread both geographically and ecologically. Moreover, because

diastatic yeasts contain in their genomes, domestication markers, not seen in wild populations of *S. cerevisiae* but present in domesticated ones (Almeida et al., 2015; Barbosa et al., 2018; Pontes et al., 2020), we posit that they derive from domesticated lineages. Our interpretation is that brewing strains have evaded the domestication niche and were able to colonize natural environments. By finding such lineages associated with fruits and traditionally fermented beverages in Botswana our best explanation is that they represent feral yeasts, i.e. formerly domesticated strains that no longer thrive in the original domestication setting where the lineage was formed. Rather than forming a subclade within the Beer 2D lineage the candidate feral strains occur intermixed with production strains. This is what would be expected if ferality is occurring repeatedly from the domesticate Beer 2D stock. However, it is important to notice that the current number of candidate feral diastatic strains is low. Further research, especially the isolation of additional diastatic strains in natural settings, is required to further assess this possibility.

Our analysis of the *STA1* gene together with its ancestor *SGA1* suggested that a limited number of gene fusion events gave rise to the extant diversity of *STA1* alleles, which likely were originated in the Wine population or in the early stages of formation of the Beer 2D population. Since our phylogeny of *STA1* representatives yielded the two main subgroups seen at the population level with whole-genome data, we conclude that *STA1* was already present when the lineage arose and therefore its divergence mimics, for most part, that of the Beer2D population.

Ferality in *Saccharomyces* has been argued previously in studies involving mostly wine yeasts (Almeida et al., 2015; Barbosa et al., 2018; De Chiara et al., 2022). However, clear evidence supporting the persistence of domesticates in natural environments substantially separated from the artificial ones that promoted domestication is scarce. For example, wine yeasts can be isolated from vineyards, which are seen by some as a natural environment. Therefore, those isolates can be considered as feral since they escaped the wine fermentation, the environment that provided the selective pressures that shaped their genotypes and phenotypes. However, given that new wine fermentations occur every year near the vineyard, these are "open" fermentations that can be colonized by the vineyard microbiota, contrary to beer brewing. Thus, it is not clear if vineyard wine yeasts are true feral yeasts that permanently abandoned the domestication environment. In this respect, the ferality of the Beer2D population documented here is substantially different, as the "closed" nature of beer fermentations, involving the use of starter cultures, does not allow for colonization from the environment during the normal fermentation process.

The lack of documented cases of ferality in the Beer 1 population, contrary to what was observed here for the Beer 2D population highlights the distinctiveness of these two groups. It appears that several attributes of the Beer 1 population, like incapacity to tolerate high temperatures or to produce ascospores renders them unfit to thrive in environments other than beer wort. On the contrary, the Beer2D yeasts, that are closer to the wine yeasts, appear to have a higher phenotypic plasticity. They can use more recalcitrant carbon sources like starch and dextrins, they are more tolerant to increased temperatures and ethanol concentrations, and they are able to form ascospores and biofilms, which can increase their resistance to environmental stresses (Meier-Dörnberg et al., 2018; Suiker et al., 2021; Suiker and Wösten 2022). The combination of all these traits supports the ferality hypothesis in this population.

Credit authors statement

Francisca Paraíso: Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing – review & editing. **Ana Pontes:** Methodology, Investigation, Formal analysis. **Joana Neves:** Methodology, Investigation. **Kebaneilwe Lebani:** Investigation, Writing– review. **Mathias Hutzler:** Methodology, Investigation Writing – review. **Nerve Zhou:** Investigation, Writing – review. **José Paulo Sampaio:** Conceptualization, Methodology, Investigation, Writing – original draft, review & editing, Supervision, Project administration, Funding acquisition.

Data availability

Sanger sequences obtained in this study are available in Supplementary Data 1 and sequence raw reads are openly available at the European Nucleotide Archive (ENA) under the BioProject ID: PRJEB59588.

Funding

This work was supported by Fundação para a Ciência e a Tecnologia (Portugal) grants UIDB/04378/2020 (UCIBIO) and LA/P/0140/2020 (i4HB).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Sérgio Romão and the staff of 8ª Colina, a Portuguese craft brewery, for their assistance during the isolation of beer yeast contaminants. This work was carried out with support of INCD funded by FCT and FEDER under project 22153-01/SAICT/2016.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fm.2023.104320.

References

Adam, A.C., Latorre-García, L., Polaina, J., 2004. Structural analysis of glucoamylase encoded by the STA1 gene of *Saccharomyces cerevisiae* (var. *diastaticus*). Yeast 21, 379–388. https://doi.org/10.1002/yea.1102.

Almeida, P., Barbosa, R., Zalar, P., Imanishi, Y., Shimizu, K., Turchetti, B., Legras, J.-L., Serra, M., Dequin, S., Couloux, A., Guy, J., Bensasson, D., Gonçalves, P., Sampaio, J. P., 2015. A population genomics insight into the Mediterranean origins of wine yeast domestication. Mol. Ecol. 24, 5412–5427. https://doi.org/10.1111/mec.13341.

Andrews, J., Gilliland, R.B., 1952. Super-attenuation of beer: a study of three organisms capable of causing abnormal attenuations. J. Inst. Brew. 58, 189–196. https://doi.org/10.1002/j.2050-0416.1952.tb02675.x.

Angebault, C., Djossou, F., Abélanet, S., Permal, E., Soltana, M. Ben, Diancourt, L., Bouchier, C., Woerther, P.L., Catzeflis, F., Andremont, A., D'enfert, C., Bougnoux, M. E., 2013. *Candida albicans* is not always the preferential yeast colonizing humans: a study in wayampi amerindians. J. Infect. Dis. 208, 1705–1716. https://doi.org/10.1093/infdis/jit389.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Barbosa, R., Pontes, A., Santos, R.O., Montandon, G.G., Ponzzes-Gomes, C.M. de, Morais, P.B., Gonçalves, P., Rosa, C.A., Sampaio, J.P., 2018. Multiple rounds of artificial selection promote microbe secondary domestication - the case of cachaça yeasts. Gen, Biol. Evol. 10, 1939–1955. https://doi.org/10.1093/gbe/evy132.

Bigey, F., Segond, D., Friedrich, A., Guezenec, S., Bourgais, A., Huyghe, L., Agier, N., Nidelet, T., Sicard, D., 2020. Evidence for two main domestication trajectories in Saccharomyces cerevisiae linked to distinct bread-making processes. Curr. Biol. https://doi.org/10.1016/j.cub.2020.11.016.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

De Chiara, M., Barré, B.P., Persson, K., Irizar, A., Vischioni, C., Khaiwal, S., Stenberg, S., Amadi, O.C., Žun, G., Doberšek, K., Taccioli, C., Schacherer, J., Petrovič, U., Warringer, J., Liti, G., 2022. Domestication reprogrammed the budding yeast life cycle. Nat. Ecol. Evol. 64 (6), 448–460. https://doi.org/10.1038/s41559-022-01671-9, 2022.

Depristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., Del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J.,

Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M. J., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–501. https://doi.org/10.1038/ng.806.

Duan, S.-F., Han, P.-J., Wang, Q.-M., Liu, W.-Q., Shi, J.-Y., Li, K., Zhang, X.-L., Bai, F.-Y., 2018. The origin and adaptive evolution of domesticated populations of yeast from Far East Asia. Nat. Commun. 9, 2690. https://doi.org/10.1038/s41467-018-05106-7.

Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A., Roncoroni, M., Voordeckers, K., Miraglia, L., Teiling, C., Steffy, B., Taylor, M., Schwartz, A., Richardson, T., White, C., Baele, G., Maere, S., Verstrepen, K.J., 2016. Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. Cell 166, 1397–1410. https://doi.org/10.1016/j.cell.2016.08.020.

Gonçalves, M., Pontes, A., Almeida, P., Barbosa, R., Serra, M., Libkind, D., Hutzler, M., Gonçalves, P., Sampaio, J.P., 2016. Distinct domestication trajectories in top-fermenting beer yeasts and wine yeasts. Curr. Biol. 26, 2750–2761. https://doi.org/10.1016/j.cub.2016.08.040.

Hutter, S., Vilella, A.J., Rozas, J., 2006. Genome-wide DNA polymorphism analyses using VariScan. BMC Bioinf. 7, 409. https://doi.org/10.1186/1471-2105-7-409.

Hutzler, M., Riedl, R., Koob, J., Jacob, F., 2012. Fermentation and spoilage yeasts and their relevance for the beverage industry - a review. BrewingScience 65, 33–52 (Not available).

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27, 722–736. https://doi.org/10.1101/GR.215087.116.

Krogerus, K., Gibson, B., 2020. A re-evaluation of diastatic *Saccharomyces cerevisiae* strains and their role in brewing. Appl. Microbiol. Biotechnol. 104, 3745–3756. https://doi.org/10.1007/S00253-020-10531-0/FIGURES/2.

Krogerus, K., Magalhães, F., Kuivanen, J., Gibson, B., 2019. A deletion in the STA1 promoter determines maltotriose and starch utilization in *STA1+ Saccharomyces cerevisiae* strains. Appl. Microbiol. Biotechnol. 103, 7597–7615. https://doi.org/10.1007/s00253-019-10021-y.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Latorre, M.A., Hutzler, M., Michel, M., Zarnkow, M., Libkind Frati, D., 2020. Genotypic diversity of *Saccharomyces cerevisiae* spoilers in a community of craft microbreweries. BrewingScience 65, 33–52. https://doi.org/10.23763/BrSc20-06latorre.

Lee, T.J., Liu, Y.C., Liu, W.A., Lin, Y.F., Lee, H.H., Ke, H.M., Huang, J.P., Lu, M.Y.J., Hsieh, C.L., Chung, K.F., Liti, G., Tsai, I.J., 2022. Extensive sampling of *Saccharomyces cerevisiae* in Taiwan reveals ecology and evolution of pre domesticated lineages. Genome Res. 32 https://doi.org/10.1101/GR.276286.121 gr.276286.121.

Letunic, I., Bork, P., 2021. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. https://doi.org/10.1093/NAR/GKAB301.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Meier-Dörnberg, T., Jacob, F., Michel, M., Hutzler, M., 2017. Incidence of *Saccharomyces cerevisiae* var. *diastaticus* in the beverage industry: cases of contamination. Master Brew. Assoc. Am. Tech. Q 54, 140–148. https://doi.org/10.1094/TQ-54-4-1130-01, 2008–2017.

Meier-Dörnberg, T., Kory, O.I., Jacob, F., Michel, M., Hutzler, M., 2018. *Saccharomyces cerevisiae* variety *diastaticus* friend or foe?—spoilage potential and brewing ability of different Saccharomyces cerevisiae variety diastaticus yeast isolates by genetic, phenotypic and physiological characterization. FEMS Yeast Res. 18, foy023. https://doi.org/10.1093/femsyr/foy023.

Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195. https://doi.org/10.1093/molbev/mst024.

Morris, S., Humphreys, D., Reynolds, D., 2006. Myth, marula, and elephant: an assessment of voluntary ethanol intoxication of the african elephant (*Loxodonta africana*) following feeding on the fruit of the marula tree (*Sclerocarya birrea*). Physiol. Biochem. Zool. 363–369. https://doi.org/10.1086/499983.

Motlhanka, K., Lebani, K., Boekhout, T., Zhou, N., 2020. Fermentative microbes of khadi, a traditional alcoholic beverage of Botswana. Fermentation 6, 51. https://doi.org/10.3390/fermentation6020051.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.

Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavory, F., Mallet, S., Cambon, B., Legras, J.-L., Wincker, P., Casaregola, S., Dequin, S., 2009. Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. Proc. Natl. Acad. Sci. U.S.A. 106, 16333–16338. https://doi.org/10.1073/pnas.0904673106.

Peter, J., De Chiara, M., Friedrich, A., Yue, J.-X., Pflieger, D., Bergström, A., Sigwalt, A., Barre, B., Freel, K., Llored, A., Cruaud, C., Labadie, K., Aury, J.-M., Istace, B., Lebrigand, K., Barbry, P., Engelen, S., Lemainque, A., Wincker, P., Liti, G., Schacherer, J., 2018. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature 556, 339–344. https://doi.org/10.1038/s41586-018-0030-5.

Pontes, A., Hutzler, M., Brito, P.H., Sampaio, J.P., 2020. Revisiting the taxonomic synonyms and populations of *Saccharomyces cerevisiae* – phylogeny, phenotypes, ecology and domestication. Microorganisms 8, 903. https://doi.org/10.3390/microorganisms8060903.

Poplin, R., Ruano-Rubio, V., DePristo, M., Fennell, T., Carneiro, M., Van der Auwera, G., Kling, D., Gauthier, L., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C., Lek, M., Gabriel, S., Daly, M., Neale, B., MacArthur, D., Banks, E., 2017. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv, 201178. https://doi.org/10.1101/201178.

Saada, O.A., Tsouris, A., Large, C., Friedrich, A., Dunham, M.J., Schacherer, J., 2022. Phased polyploid genomes provide deeper insight into the multiple origins of domesticated *Saccharomyces cerevisiae* beer yeasts. Curr. Biol. 32, 1350–1361.e3. https://doi.org/10.1016/J.CUB.2022.01.068.

Suiker, I.M., Arkesteijn, G.J.A., Zeegers, P.J., Wösten, H.A.B., 2021. Presence of *Saccharomyces cerevisiae* subsp. *diastaticus* in industry and nature and spoilage capacity of its vegetative cells and ascospores. Int. J. Food Microbiol. 347, 109173 https://doi.org/10.1016/J.IJFOODMICRO.2021.109173.

Suiker, I.M., Wösten, H.A., 2022. Spoilage yeasts in beer and beer products. Curr. Opin. Food Sci. 44, 100815 https://doi.org/10.1016/J.COFS.2022.100815.

Wang, Q.-M., Liu, W.-Q., Liti, G., Wang, S.-A., Bai, F.-Y., 2012. Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. Mol. Ecol. 21, 5404–5417. https://doi.org/10.1111/j.1365-294X.2012.05732.x.

Will, J.L., Kim, H.S., Clarke, J., Painter, J.C., Fay, J.C., Gasch, A.P., 2010. Incipient balancing selection through adaptive loss of aquaporins in natural *Saccharomyces cerevisiae* populations. PLoS Genet. 6, e1000893 https://doi.org/10.1371/journal.pgen.1000893.

Yamashita, I., Nakamura, M., Fukui, S., 1987. Gene fusion is a possible mechanism underlying the evolution of *STA1*. J. Bacteriol. 169, 2142–2149. https://doi.org/10.1128/jb.169.5.2142-2149.1987.