environmental microbiology

Environmental Microbiology (2021) **23**(6), 3182–3195 doi:10.1111/1462-2920.15540

Global distribution of *IRC7* alleles in *Saccharomyces cerevisiae* populations: a genomic and phenotypic survey within the wine clade

Javier Ruiz,[1] **Miguel de Celis**,[1] **María Martín-Santamaría**,[1] **Iván Benito-Vázquez**,[1] **Ana Pontes**,[2] **Val F. Lanza**,[3] **José Paulo Sampaio**,[2] **Antonio Santos**[1*] **and Ignacio Belda**[1*]

[1]*Department of Genetics, Physiology and Microbiology. Unit of Microbiology. Biology Faculty, Complutense University of Madrid, Madrid, 28040, Spain.*

[2]*Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, 2829-516, Portugal.*

[3]*Department of Microbiology, Ramón y Cajal University Hospital, IRYCIS, Madrid, 28034, Spain.*

Summary

The adaptation to the different biotic and abiotic factors of wine fermentation has led to the accumulation of numerous genomic hallmarks in *Saccharomyces cerevisiae* wine strains. *IRC7*, a gene encoding a cysteine-S-β-lyase enzyme related volatile thiols production in wines, has two alleles: a full-length allele (*IRC7*^{F}) and a mutated one (*IRC7*^{S}), harbouring a 38 bp-deletion. Interestingly, *IRC7*^{S}-encoding a less active enzyme – appears widespread amongst wine populations. Studying the global distribution of the *IRC7*^{S} allele in different yeast lineages, we confirmed its high prevalence in the Wine clade and demonstrated a minority presence in other domesticated clades (Wine-PDM, Beer and Bread) while it is completely missing in wild clades. Here, we show that *IRC7*^{S}-homozygous (HS) strains exhibited both fitness and competitive advantages compared with *IRC7*^{F}-homozygous (HF) strains. There are some pieces of evidence of the direct contribution of the *IRC7*^{S} allele to the outstanding behaviour of HS strains (i.e., improved response to oxidative stress conditions and higher tolerance to high copper levels); however, we also identified a set of sequence variants with significant co-occurrence patterns with the *IRC7*^{S} allele, which can be co-contributing to the fitness and competitive advantages of HS strains in wine fermentations.

Received 24 September, 2020; revised 23 March, 2021; accepted 19 April, 2021. *For correspondence. E-mail ansantos@ucm.es; Tel. (+34) 913944962. E-mail ignaciobelda@ucm.es; Tel. (+34) 913944962; Fax: (+34) 913944964.

Introduction

Saccharomyces cerevisiae is a eukaryotic model microorganism used in cellular physiology, molecular biology and genetics. However, much is still unknown about its metabolism in natural environments, beyond the laboratory (Liti, 2015). *S. cerevisiae* strains are worldwide distributed, occurring in multiple wild habitats and associated to several human activities (Wang *et al*., 2012). From its natural origin, the genome and the phenome of *S. cerevisiae* strains have been shaped for millennia, through different domestication events driven by a combination of natural and anthropic selection forces. These domestication events are accompanied with specific phenotypic traits originating from genetic variants, from single-nucleotide polymorphism to copy number variation or horizontal gene transfer (Belda *et al*., 2020). This process has originated highly adapted strains, defining different – wild and domesticated – lineages of *S. cerevisiae* during its evolutionary history (*domesticated clades*: Wine, Wine-PDM (Prise de Mousse), Beer1, Beer2, Bread and Sake; *wild clades*: West African, Mediterranean Oak, North America & Japan and Malaysia; Gallone *et al*., 2016; Gonçalves *et al*., 2016). These clades reflect, not only their geographical distribution, but also their lifestyle in association to different niches (Liti *et al*., 2009; Schacherer *et al*., 2009; Strope *et al*., 2015; Barbosa *et al*., 2018; Peter *et al*., 2018; Pontes *et al*., 2020). Because of the selective pressures of the different niches and, as a consequence of their high genome plasticity, *S. cerevisiae* strains are highly adapted to each environment they inhabit (Legras *et al*., 2018).

One of these well-defined monophyletic lineages is the Wine clade, including isolates from wine must, grapes and vineyard soils (Almeida *et al*., 2017). It has been demonstrated that this clade arose from a single bottleneck event of domestication (Peter *et al*., 2018),

Mediterranean Oak population being the proposed wild origin for wine yeasts clade (Almeida *et al*., 2015). Thereafter, viticulture expansion through Europe and Mediterranean Sea led to the migration of yeasts associated to grapevines to all the wine-producing regions around the world (Legras *et al*., 2007). For approximately 9000 years (McGovern *et al*., 2004), the specific environmental conditions of wine fermentations and the human pressure to achieve wine desirable traits, have led to a rapid evolutionary adaptation of wine strains, implying important changes on their genomes if compared with non-wine strains (Marsit and Dequin, 2015). Therefore, numerous hallmarks of domestication have been reported in wine strains, as examples of the adaptive process to wine environment (Belda *et al*., 2020).

In this work, we report new genetic and phenotypic signatures within the wine yeasts population which help to explain the previously reported paradoxical distribution of the *IRC7*^{S} allele in wine yeasts (Belda *et al*., 2016). This gene, encoding a cysteine-S-conjugate β-lyase (EC 4.4.1.13), is the direct responsible for the release of volatile thiols (mainly 4MSP (4-methyl-4-sulfanylpentan-2-one)) from their cysteinylated precursors, conferring pleasant aromas in white wines (Harsch and Gardner, 2013; Howell *et al*., 2004; Swiegers and Pretorius, 2007; Tominaga *et al*., 1998). Two alleles have been reported for *IRC7* in *S. cerevisiae*: a 1200-bp full-length allele (*IRC7*^{F}) encoding a 400 amino acid protein, and an altered allele harbouring a 38-bp deletion (*IRC7*^{S}), creating a premature stop codon, and thus encoding a shorter enzyme (360 amino acids) with lower activity (Roncoroni *et al*., 2011). Hence, three *IRC7* genotypes have been described in *S. cerevisiae* strains: homozygous strains for the full-length *IRC7* allele, *IRC7*^{F} (HF), heterozygous strains (HT), and homozygous strains for the short-length *IRC7* allele, *IRC7*^{S} (HS). Surprisingly, the great majority of *S. cerevisiae* wine strains are homozygous for the *IRC7*^{s} allele (HS strains), and therefore, they encode a less active β-lyase enzyme to release aromatic thiols (Roncoroni *et al*., 2011; Belda *et al*., 2016; Cordente *et al*., 2019).

To understand the high prevalence of HS strains in wine environment, we performed: (i) an *IRC7* genotyping survey in a global collection of *S. cerevisiae* genomes coming from different domesticated and wild lineages; (ii) a genome-wide association study of *IRC7*-related sequence variants found in genes potentially involved the population prevalence of strains harbouring *IRC7*^{S} allele in domesticated populations and (iii) a high throughput phenotyping study to seek metabolic and growth differences between the three *IRC7* genotypes in wine strains. In summary, here we evidence that *IRC7*^{S}-harbouring strains exhibited fitness and competitive advantages which could explain the outstanding high prevalence of these strains in wine population. Also, we found some phenotypic patterns in HS strains (i.e., an overall improvement in proliferative growth parameters, a better competitive fitness against HF strains and a strong pattern of resistance to *S. cerevisiae* killer strains) that seems to go beyond the functional activity of Irc7p. Thus, here we reported a set of sequence variants, with a high co-occurrence pattern with the *IRC7*^{S} allele which could help to explain the population prevalence, and the fitness and competitive advantages of HS strains in wine fermentations.

Results

IRC7 allele distribution in S. cerevisiae *populations*

The evolutionary history of *S. cerevisiae* has defined well-separated domesticated and wild phylogenetic clades (Peter *et al*., 2018). To study the prevalence of the *IRC7* alleles amongst *S. cerevisiae* populations, we performed an *IRC7* genotyping survey using a collection of 283 *S. cerevisiae* genomes representing different – domesticated and wild – phylogenetic clades and origins (Supporting Information File S1). As stated before, two *IRC7* alleles have been described (*IRC7*^{F} and *IRC7*^{S}), defining three genotype groups: HF for homozygous strains for the *IRC7*^{F} allele; HT for heterozygous strains; and HS for homozygous strains for the *IRC7*^{S} allele (Roncoroni *et al*., 2011).

Figure 1 shows the distribution of *IRC7* alleles amongst different *S. cerevisiae* phylogenetic clades. Here we confirmed that the *IRC7*^{S} allele was largely widespread in the Wine clade (85% of the strains). However, other domesticated clades such as Wine-PDM (‘*Prise de Mousse*’, a cluster of strains isolated from the secondary fermentation of sparkling wines), Beer2 and Bread showed a minority presence of the *IRC7*^{S} allele (27.3%, 35.0% and 37.5% respectively). All the strains pertaining to the other domesticated (Beer 1 and Sake), and wild clades (Mediterranean Oak, Philippines, North America & Japan, West Africa and Malaysia) only harboured the *IRC7*^{F} allele.

IRC7-rooted genomic survey in domesticated populations

With the aim of understanding if some additional genomic features – beyond the *IRC7*^{S} allele – can be co-contributing to the population prevalence of HS strains in domesticated populations, especially amongst wine strains, we carried out a genomic survey looking for allelic variants potentially associated with the *IRC7*^{S} allele. We performed a preliminary search in the whole genome of 9 wine strains (HS4, HS6, HS9; HT3, HT6, HT10;

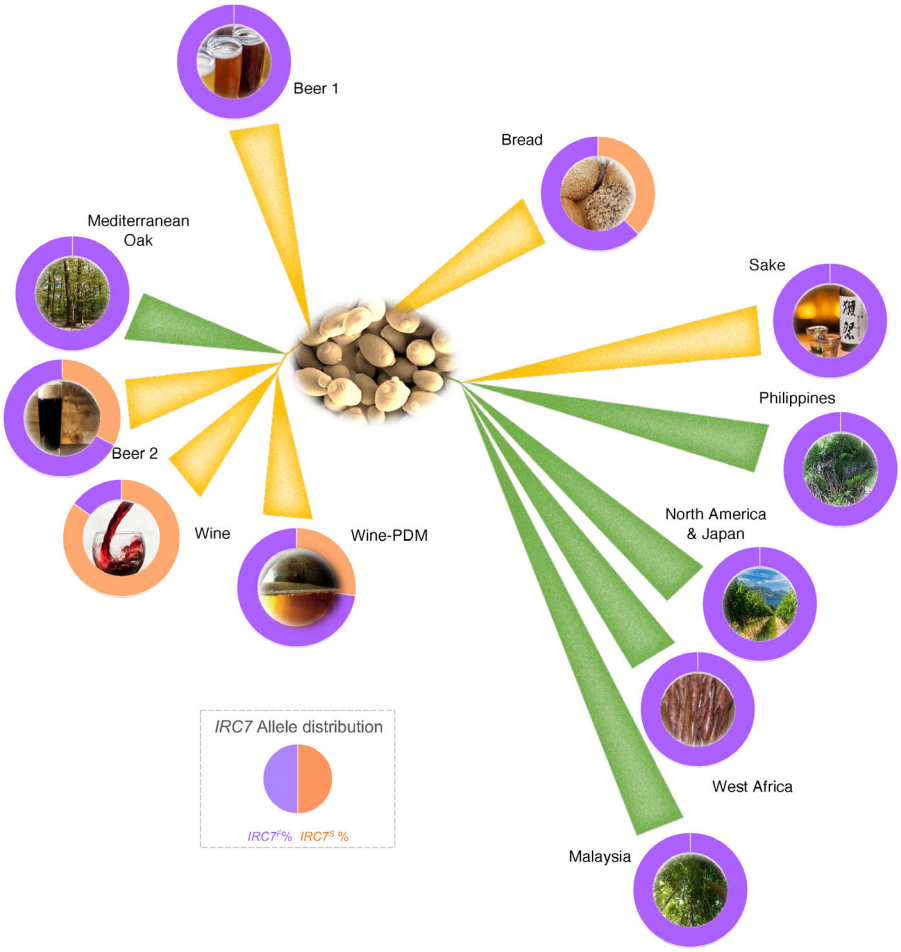

Fig. 1. Schematic overview of the genetic relationships amongst *S. cerevisiae* lineages (phylogeny inferred using maximum likelihood) and prevalence of the *IRC7* alleles (orange for $IRC7^{S}$ and purple for $IRC7^{F}$) on each phylogenetic clade. Yellow branches highlight domesticated clades and green branches highlight wild clades.

HF1, HF2, HF9; see Supporting Information File S1), including 3 representative strains of each *IRC7* genotype that were selected amongst those that will be characterized phenotypically later on (Table S2). In this search, we used *S. cerevisiae* VL3 (a well-studied and widely used strain in the wine industry; homozygous for the $IRC7^{F}$ allele) as the reference genome for the variant calling. The presence and co-occurrence of these sequence variants amongst the nine studied strains were represented in a bipartite network, including moderate and highly important mutations, and discarding those widespread mutations found in all the strains and those rare ones just found in one single strain (Fig. S1). Interestingly, HS strains appeared closely located in the network, indicating a conserved pattern of sequence variants against the HF strain genome (VL3) used as a reference. To focus our analysis on the potential genomic features associated with the $IRC7^{S}$ allele, we identified those genetic variants found as conserved amongst the genome of all the HS strains and absent in all HF strains in the subset of nine genomes. With this premise, the three HS strains shared 34 sequence variants – pertaining to 10 different genes, including *IRC7* – which were not detected in any of the three HF strains analysed here (Table S1).

The 34 sequence variants, identified in the preliminary search, were then explored in a wider population survey of 150 genomes from different domesticated clades (genomes highlighted with an asterisk in Supporting Information File S1; mainly obtained from Gallone and colleagues (2016) and the genomes from the sequencing effort of this work), to confirm those sequence variants showing significant co-occurring patterns with the $IRC7^{S}$ allele (Fig. 2A, Supporting Information File S2). Figure 2B shows the prevalence patterns of the most significant sequence variants in terms of co-occurrence rates with $IRC7^{S}$, within the wine clade and in the other domesticated clades (see Table S1 for a description of the nature and effect of the sequence variant detected, and the function of the genes affected). As occurred with $IRC7^{S}$, all these sequence variants showed higher prevalence figures within the wine clade than in other domesticated clades (with the exception of XII_2022 *PAU6*) (Fig. 2B). Particularly noticeable, some sequence variants showed extremely low (those found in *SDH1, YDR185C* and

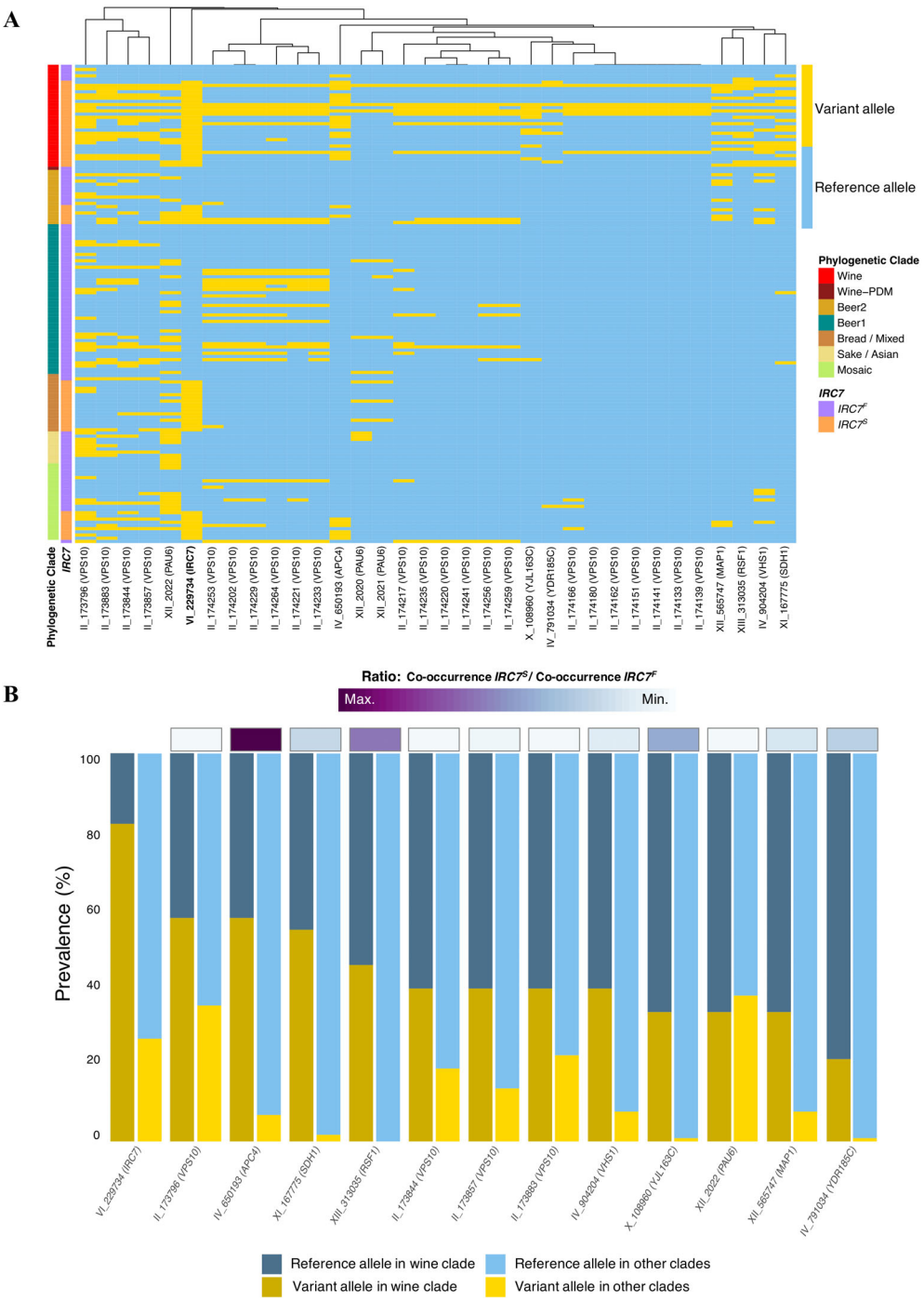

Fig. 2. A. Presence of the 34 sequence variants (affecting to 10 different genes), identified in the variant calling analysis and potentially associated with *IRC7*^{S} allele, in 150 genomes from different domesticated clades (Wine, Wine-PDM, Sake/Asian, Bread/Mixed, Beer1, Beer2 and Mosaic). The *IRC7* allele and phylogenetic clades of the strains are highlighted at the left side of the heatmap, and the sequence variant of *IRC7* is highlighted in bold letters in the heatmap. The presence/absence (variant allele/reference allele) of each sequence variant in the genome collection (representing six domesticated clades) is indicated by yellow and blue colours respectively. Strains (represented as rows in the heatmap) are grouped by their phylogenetic clade and *IRC7* genotype. Sequence variants (represented as columns in the heatmap) are clustered in the dendrogram based on their distribution amongst genomes. Supporting data for this heatmap is described in the Supporting Information File S2.
B. Prevalence of the allelic variants amongst yeast strains belonging to the wine clade and average values for the other non-wine domesticated clades. Only those sequence variants with a prevalence in Wine clade higher than 25% were represented. The global co-occurrence ratios between each allelic variant with the *IRC7*^{S} and *IRC7*^{F} alleles were also represented in the heatmap above the barplot.

YJL163C genes) to no prevalence (that found in *RSF1*) in non-wine domesticated clades. None of these sequence variants has higher prevalence figures than *IRC7*^{S} within the wine clade, but one of the variants found in *VPS10* (II_173796) and the ones found in *APC4* and *SDH1* have a prevalence higher than 50% amongst wine strains. In addition, the sequence variants found in *APC4* and *RSF1* stand out as those with the highest co-occurrence ratio with *IRC7*^{S} versus *IRC7*^{F} allele (Fig. 2B, Supporting Information File S3).

Phenotypic landscape in wine strains based on IRC7-genotype

To understand the biological basis of the high prevalence of the *IRC7*^{s} allele within the Wine clade, and due to the great diversity of functions in which the genes carrying mutations potentially associated with *IRC7*^{S} are involved, we carried out a high throughput phenotyping looking for differences in growth parameters between the *IRC7*-genotype groups. Thirty *S. cerevisiae* wine strains, representing the three different *IRC7* genotypes (10 HF, 10 HT and 10 HS; Table S2) were assayed in a panel of 48 culture conditions (testing different carbon and nitrogen sources, physicochemical conditions and antimicrobials; Table S3). Growth curves in axenic cultures were analysed to obtain the following growth parameters: lag time (time to initiate proliferation), growth rate (during exponential growth) and proliferative efficiency (population density change; Supporting Information File S4).

Figure 3A represents mean values (normalized to average for each condition) of the growth parameters obtained in all the growth conditions tested for the group of strains pertaining to each *IRC7* genotype. Despite the high variability observed in lag phase duration and growth rate between strains – especially within the HF group-, HF strains showed an overall lower fitness than HS strains, exhibiting, on average, longer lag phases and lower growth rates ($P < 0.05$) in most conditions tested (Fig. 3A, Fig. S2). Conversely, we did not find significant differences on the proliferative efficiency values amongst the *IRC7*-genotype groups. In addition, we observed clustered growth patterns highly associated with the *IRC7*-genotype groups (Fig. S2), where HS strains exhibited an increased overall fitness.

In addition, metabolite profiles in wine fermentations can display other subtle differences between strains with different genotypes (Legras *et al*., 2018). Thus, we performed microvinifications in synthetic grape must to evaluate strain-associated patterns in basic physicochemical wine parameters. Despite the notable variability amongst the strains belonging to each *IRC7* genotype groups (Table S4), Fig. 3B proves that the global metabolite profile of HF and HS strains defines significantly different clusters, while no significant differences were obtained between HS and HT groups (PERMANOVA, $P < 0.01$). As the PC1 (accounting for a 30.8% of the explained variance) appeared as the main component separating HS and HF metabolite clusters (Fig. 3B), we should highlight the higher levels of ethanol and the lower pH showed by HS strains (Table S4) as the main drivers of the wine metabolite clustering.

Competitive ability of wine strains based on IRC7-genotype

To evaluate the potential advantage of HS strains to colonize wine fermentations, we carried out yeast-to-yeast (HS vs. HF) competition experiments under fermentative (wine and beer fermentations) and non-fermentative (glycerol-based medium) conditions. We performed these competition assays with three strains representing the HS genotype and four strains representing the HF genotype, where HS and HF strains were co-inoculated at the same (1:1) proportion The implantation of each strain (percentage of the final population represented by each strain) in the competition experiments were determined by interdelta polymorphisms genotyping of the colonies obtained at the end the competition assays. Figure 4A shows a better competitive performance of HS strains, with mean implantation percentages of 65%–82%–75% and success rates of 67%–100%–83% in wine, beer and non-fermentative conditions respectively.

Looking for other pieces of evidence, explaining the competitive advantage showed by *IRC7*-HS strains, we also evaluated the killer phenotype in our strains collection, looking for patterns in killer activity and/or resistance amongst HS–HT–HF groups. Figure 4B shows the results obtained for the killer (growth inhibition) assays, representing the sensitivity/resistance phenotypes of the strains tested. Killer phenotype showed a pattern associated with the *IRC7* genotype, as all HS strains tested were resistant to the rest of *S. cerevisiae* strains tested as potential killer strains. Conversely, 7 out of 10 HT-strains and only 2 out of 10 HF strains were resistant. In addition, killer activity (against at least one of the potentially sensitive strains tested) was observed in 9 out of 10 HS strains, 7 out of 10 HT-strains, and only in 4 out of 10 HF strain.

Discussion

As reported before (Roncoroni *et al*., 2011; Belda *et al*., 2016; Cordente *et al*., 2019), the *IRC7*^{S} allele appeared widespread in yeasts within the wine clade. Interestingly, here we show that it is also present, with a minority prevalence, in other domesticated clades such as Bread and, Wine-PDM and Beer2 – both defined as intermediate groups between non-wine and wine strains

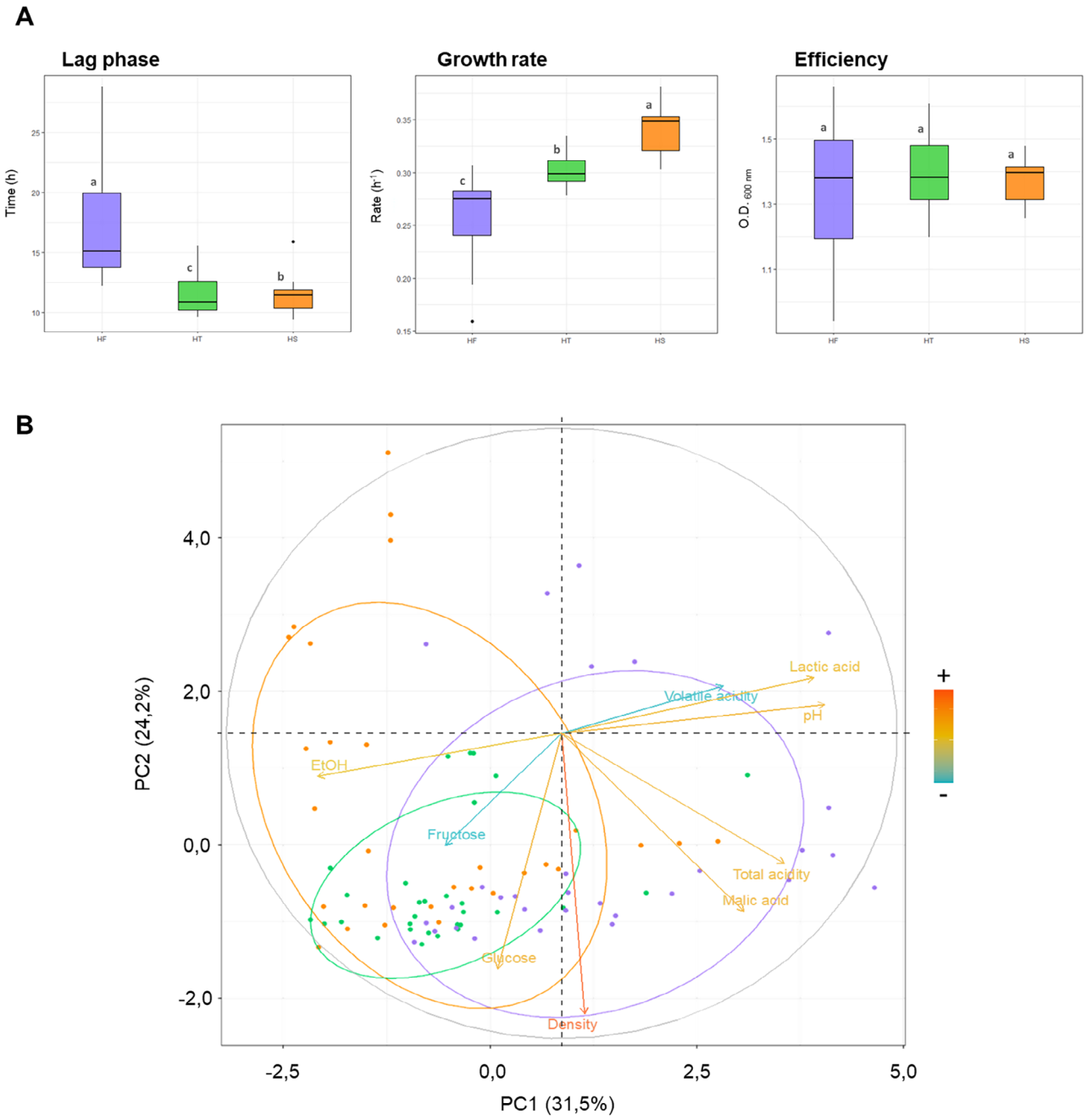

Fig. 3. A. Boxplot of the average lag time, growth rate and efficiency of the three different strains groups, according to their *IRC7* genotype. The performance in all the media panel were considered together in each plot. Different letters indicate the existence of statistical differences ($P < 0.05$).
B. Principal component analysis (PCA) representation of the wine fermentative parameters measured after the fermentation of the strains in synthetic grape must. Pairwise comparisons using PERMANOVAs on the Euclidean distance matrix used to plot the PCA reported significant differences between HF-HS ($P = 0.003$) and HT-HF ($P = 0.024$), but not between HS-HT ($P = 0.171$). [Color figure can be viewed at wileyonlinelibrary.com]

(Dunn *et al*., 2012; Gallone *et al*., 2016)-, but it is completely absent in wild clades. In addition, we detected the *IRC7*^{S} allele in wine strains isolated from all continents (Europe, America, Asia, Africa and Oceania); thus, the occurrence of this allele does not appear as restricted to a specific region (Supporting Information File S1). All these observations suggest that the prevalence of the *IRC7*^{S} allele may have been favoured during the domestication process of *S. cerevisiae*, with an outstanding prevalence in wine strains.

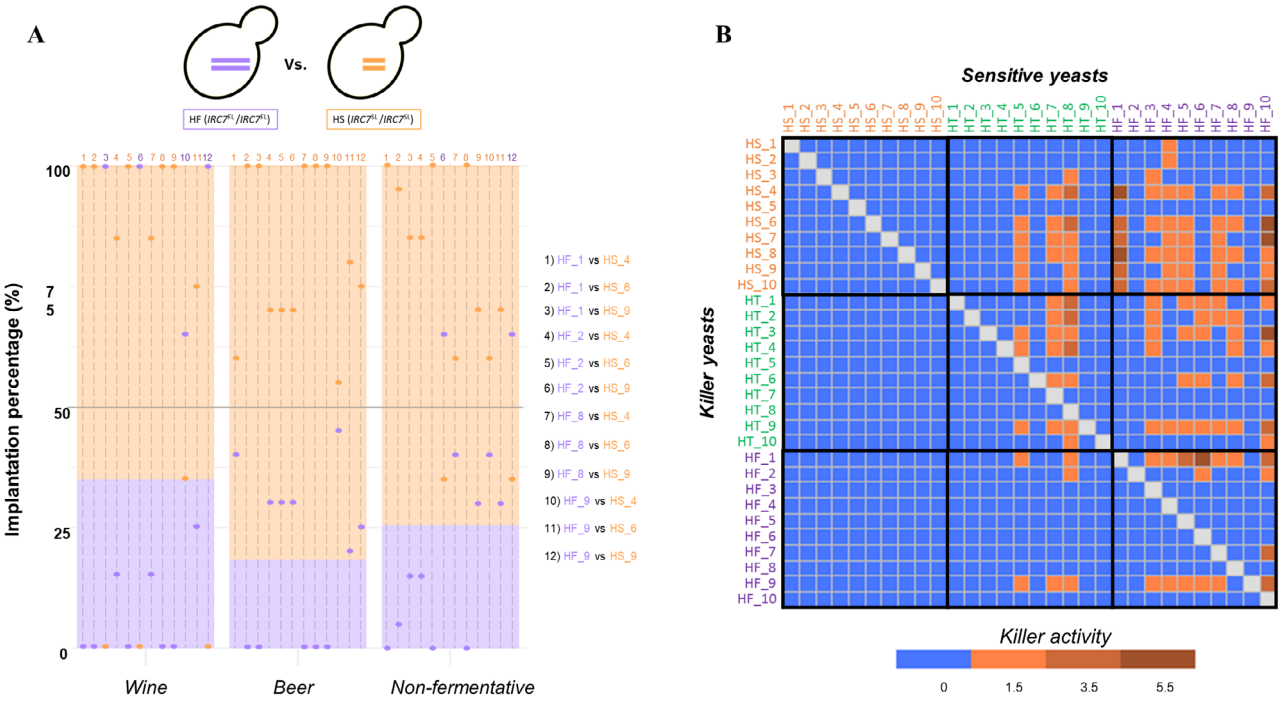

Fig. 4. A. Competition assays of three HS strains (HS4, HS6, HS9) and four HF strains (HF1, HF2, HF8, HF9) in wine, beer and non-fermentative mimic (glycerol-based) media. Twelve combinations of HS and HF strains (1:1 proportion) were assayed in the three different conditions. Implantation percentage after 7 days of culture, analysed by interdelta polymorphism fingerprinting, is represented. Bar plots indicate the average implantation of each *IRC7* group and dot plots indicate the implantation percentage of each competition assay. A global comparison of the implantation of HS and HF strains in all growth conditions showed significant differences (*T*-test, $P < 0.001$) with higher figures for HS strains.
B. Killer activity of the yeast crosses assays of 30 strains representing each *IRC7*-genotype group (HS, orange, HT, green and HF purple) is represented. Killer yeasts term refers to the yeast tested for killer activity and sensitive yeasts refers to yeast tested for killer sensitivity. Orange colour indicates killer activity and blue colour indicates the absence of any growth inhibition. Semiquantitative data on the intensity of the killer activity – calculated as the total diameter of the halo (growth inhibition) divided by the diameter of the biomass concentrated zone – is represented by different orange shades.

Reinforcing this hypothesis, other wine yeasts not belonging to the *Saccharomyces* genus, such as *Torulaspora delbrueckii* (a yeast species with a high capacity to release volatile thiols in wine fermentations), do not harbour the $IRC7^S$ allele in the orthologous gene in any of the wine strains investigated, but only the $IRC7^F$ allele (Belda *et al*., 2017). In this sense, the prevalence of the $IRC7^S$ allele in *S. cerevisiae* wine strains may be counterintuitive, as the $IRC7^F$ allele is strongly related with the production of pleasant aromas in wine fermentations and, as occurred with other aroma-related genes, those alleles related with higher quality products are assumed to be artificially selected through the domestication process (Steensels *et al*., 2019). Thus, in this work we aim to contribute to the pending task of understanding the biological basis of the paradoxical distribution of *IRC7* alleles in wine yeasts (Santiago and Gardner, 2015; Belda *et al*., 2016; Cordente *et al*., 2019).

Irc7p may have an implication in *S. cerevisiae* growth under some specific conditions, due to their role in cysteine and methionine metabolisms. Santiago and Gardner (2015) proposed a putative role of $IRC7^F$ coding enzyme on cysteine homeostasis by demonstrating its cysteine desulfhydrase activity. A fully functional Irc7p (encoded by the $IRC7^F$ allele) could compromise the availability of the intracellular cysteine pool, therefore, due to the important impact of cysteine on glutathione production, *IRC7* could have a relevant role in the protection against oxidative damage in fermentation environments (García-Ríos and Guillamón, 2019). This hypothesis was supported by the fact that we found slightly lower oxidative damage levels in HS strains compared with HF strains in a controlled oxidative-stress shock assay (Fig. S3A). In addition, *IRC7* is regulated by copper availability (*IRC7* expression is inhibited under high copper conditions) and it may have a role on copper tolerance (Cordente *et al*., 2019). This trait represents a competitive advantage for wine strains, since high levels of copper in grape musts (mainly coming from the use of phytosanitary sprays in the vineyard) can be toxic for yeasts, causing stuck fermentations (Cavazza *et al*., 2013). Cordente and colleagues (2019) proposed that an Irc7p enzyme with a reduced ability to utilize cysteine ensure a higher intracellular cysteine pool, needed

for the synthesis of cysteine-rich copper metallothioneins (i.e., Cup1p) which confer copper resistance to wine yeasts (Crosato *et al*., 2020). Reinforcing this hypothesis, Fig. S3B shows that HS strains have an overall better growth performance in the copper supplemented medium, exhibiting, on average, outstanding shorter lag phases and higher growth rates compared with HF strains. At this point, we should highlight the sequence variant detected in *MAP1* (significantly co-occurring with the *IRC7*^{S} allele, and showing a higher prevalence within the wine clade than in other domesticated clades; Fig. 2B), due to its functional relationship with Irc7p, as both participate in the metabolism of methionine, which has a great importance in both fermentation kinetics and wine flavour (Gutiérrez *et al*., 2013).

Apart from these observations, the other phenotypic results presented here – and not directly related to Irc7p enzymatic activity – could demonstrate an advantage phenotype of HS strains, explaining their high prevalence within Wine clade. When inoculated in grape must, yeast cells have to adapt to a highly stressful and competitive environment, performing a switch between respiration and fermentation metabolism, which is the main factor determining lag phase duration (Vermeersch *et al*., 2019). Therefore, the rapid adaptation to these conditions, which leads to the onset of exponential growth, will be decisive for the subsequent population to compete efficiently with the microbial populations inhabiting the same niche. Shorter lag phases and higher growth rates are major determinants of competitive fitness in complex multi-strains and multi-species environments (Schmidt *et al*., 2020), so, the great dominance of HS strains in wine fermentations (Fig. 4A) can be partially supported by their advantageous proliferation kinetics (Fig. 3A). However, the great diversity of phenotypes observed here, that goes beyond the cysteine-S-conjugate β-lyase activity of Irc7p, led us to hypothesize that the additional genetic variants detected by the genomic survey in other genes, but highly co-occurring with the *IRC7*^{S} allele, could help explain the population prevalence and the phenotypic pattern of HS strains.

Amongst the sequence variants identified as highly co-occurring with the *IRC7*^{S} allele (Supporting Information File S3; Fig. 2B), and therefore, potentially candidates to explain the outstanding phenotype of HS strains, *SDH1* is a gene involved in the yeast's respiratory metabolism, encoding a flavoprotein subunit of the succinate dehydrogenase (SDH), which oxidizes the succinate in the TCA cycle. The sequence variant detected in *SDH1* appeared widespread within the wine clade, but with a minority prevalence in other domesticated clades. It has been reported that, during fermentation, TCA cycle is interrupted at the level of SDH complex, but succinate is still formed by the oxidative branch of the TCA pathway (Camarasa *et al*., 2003). The transition to a respiratory metabolism and the maintenance of certain degree of respiration during fermentation have a direct impact on ethanol consumption at the end of the fermentation, and therefore, on the competitive performance during wine fermentation (Gasmi *et al*., 2014). Further studies are necessary to understand the functional impact of the sequence variant detected here in *SDH1* (genomic position: XI_167775) in the metabolism of succinate and the maintenance of respiratory activity during wine fermentation. In addition, the sequence variants found in two other genes (*VHS1* and *APC4*), should be also studied in detail to understand their potential contribution to the fitness advantage found in HS strains. Briefly, *VHS1* encodes for a protein kinase activated by glucose availability, and it has been described as a member of the fermentome group (genes required to accomplish wine fermentation in *S. cerevisiae*; Walker *et al*., 2014), and *APC4* encodes for a ubiquitin ligase involved in the anaphase inhibitors degradation and the reduction on its functionality could generate an increased competitive fitness (Breslow *et al*., 2008).

RSF1, another gene with a sequence variant highly co-occurring with the *IRC7*^{S} allele, encodes a transcriptional factor required for the transition to respiratory growth. It is specifically necessary for the use of glycerol and ethanol as carbon sources (Lu *et al*., 2005), but it is also involved in the sporulation process. Interestingly, Gerke and colleagues (2009) reported that the sequence variant in *RSF1* found in our work (XIII_93636_C, D181G) is missing in wild strains (i.e. oak strains) exhibiting a high-sporulation efficiency, but occurs in most strains isolated from vineyards. Our results reinforced this observation, showing that this *RSF1* mutation is only found in *S. cerevisiae* wine strains and completely missing in any other domesticated clades (Fig. 2B, Supporting Information File S2). In *S. cerevisiae* wild strains, sexual reproduction – and therefore sporulation lifestyle – is favoured, rather than domesticated strains which present a predominant asexual lifecycle (Liti, 2015). Thus, although further studies are necessary to demonstrate the increased ability of *RSF1*-mutated strains to survive and compete in wine fermentations, our results suggest that this sequence variant has been favoured through the domestication process of wine strains. Interestingly, as sporulation and pseudohyphal growth are conflicting behaviours (Cullen and Sprague, 2012), and the latter gives cells an advantage in food foraging at nutrient-limited conditions (Carstens *et al*., 1998), our results shows that, while a 100% of our HS strains tested showed a notable pseudohyphal growth ability when nitrogen nutrient limitations is reached, only a 30% of HF strains showed this phenotype (Fig. S4).

Furthermore, *S. cerevisiae* strains able to secrete killer toxins – active against other sensitive yeasts, including strains of the same species, but also those strains resistant to the toxins produced by killer strains, could have a significant advantage in competitive environments such as wine fermentation (Marquina *et al*., 2002). In this sense, Servien□ and colleagues (2012) demonstrated that the deletion of *VPS10* led to a K2 killer toxin hypersensitivity caused by a defect in cell osmoregulation. Thus, we suggest that the mutated version detected in *VPS10* (which accumulate several sequence variants with a higher prevalence in *IRC7*-HS than in *IRC7*-HF strains) could play a role in the strong resistance pattern to killer toxins found in most HS strains (Fig. 4B). In addition, although the biological role of *PAU6* is still unknown, the high percentage identity (84%) between Pau6p and Pau5p – which has a direct role in the resistance against yeast killer toxins (Rivero *et al*., 2015) – allows us to hypothesize that *PAU6*, and the sequence variants found on it (genomic positions: XII_2020 to XII_2022), may also be involved in killer toxins resistance, both in wine and non-wine strains.

At this point, it is important to note that, grouping the strains studied in the proliferative phenotyping (Supporting Information File S4) by all the potential genotypes (homozygous for the reference allele, heterozygous, and homozygous for the variant allele) defined by the existing allelic variants of the genes discussed above, we confirmed that the different *IRC7* genotypes exhibited the most significant differences in terms of proliferative fitness (lag time and growth rate). However, the homozygous genotypes for the allelic variants found in *VHS1, YJL163C, APC4, SDH1* and *RSF1* also showed an increased growth rate compared with the respective homozygous genotypes for the reference allele (Table S5); always with a lower statistical significance than the *IRC7* genotypes, and without any significant impact in the lag phase duration (with the exception of *YJL163C*).

Some of the results discussed in this work may partially explain the great prevalence of HS strains in the Wine clade, but the reason why the $IRC7^S$ allele is not as widespread in other highly domesticated clades like Wine-PDM, Beer or Bread (Fig. 1), remains to be unravelled. One explanation is that these clades, in comparison to Wine, are inhabiting fermentation processes from raw materials with a lower microbial diversity – sometimes even sterilized – where the strains used are not exposed to such highly competitive pressure as it occurs in grape musts (Conacher *et al*., 2019).

To conclude, since *IRC7* genotype, itself, does not seems to justify all the aspects of the advantageous fitness and competitive capability of HS strains, here we identified a set of sequence variants with high co-occurrence rates with the $IRC7^S$ allele, which may be co-contributing to the outstanding phenotype of HS strains and the population prevalence in wine environments. However, further molecular studies are needed to unravel the actual individual or joint contribution of the identified sequence variants, and to explore other complex mechanisms, such as genetic hitchhiking, which may be responsible for the phenotypic advantages found in *IRC7*-HS strains.

Experimental procedures

Yeast strains

Saccharomyces cerevisiae strains used in this study (Table S1) were from CYC (Complutense Yeast Collection, Madrid, Spain) and Agrovin S.A. (Alcázar de San Juan, Spain). Sabouraud medium (Oxoid, Hampshire, UK) was routinely used for handling of the strains.

IRC7 genotyping

Previously, our research group performed the *IRC7* genotyping by PCR analysis (Roncoroni *et al*., 2011) of a vast collection of *S. cerevisiae* wine strains (Belda *et al*., 2016). Continuing this work, we performed an extensive *IRC7* genotyping using the genome information of 283 *S. cerevisiae* genomes, representing different phylogenetic clades (Supporting Information File S1). A local BLAST database was set up for each genome and, using BLASTN searches (1e-4 *E-value* cut-off), we identified the presence of each allelic variant of *IRC7* in the genome collection, using the $IRC7^F$ sequence of *S. cerevisiae* VL3 strain as a query (GenBank: CM001131.1:228715-229917). When the BLAST search found the $IRC7^S$ allele as the subject, we then identified this genome as harbouring $IRC7^S$; thus, heterozygosis was not considered in this search. Strains harbouring $IRC7^F$ allele showed ≥99.9% of query coverage, while stains harbouring the $IRC7^S$ allele presented 98.5% of query coverage. This identity percentage difference corresponds to the 38-bp deletion fragment. The prevalence of the *IRC7* alleles amongst *S. cerevisiae* populations was represented in a schematic overview of a phylogenetic tree of both wild and domesticated of *S. cerevisiae* (Fig. 1). This tree was drawn based on a phylogenetic analysis (not included) performed with a reduced dataset of *S. cerevisiae* genomes, representing different lineages.

IRC7-phenotyping study

High throughput phenotyping. We developed a phenotyping screening to characterize the fitness of the

strains, in order to find growth ability patterns associated with the *IRC7* genotype. Thirty *S. cerevisiae* strains (Table S2; 10 strains belonging to each *IRC7-genotype* group as described lately) were subjected to a high throughput phenotyping study. Strains were precultured during 48 h in 300 µl of Synthetic Defined (SD) medium (Warringer *et al*., 2011) with some modifications (2% glucose, 0.14% Yeast Nitrogen Base (BD Difco™, USA), 0.5% ammonium sulfate, 2.27% succinic acid disodium salt and 0.077% Complete Supplement Mixture (CSM, MP Biomedicals™, UK); pH was set to 5.8). Thereafter, strains were inoculated at a final $OD_{600\ nm}$ of 0.2 into 300 µl of the specific SD medium. Assays were performed in triplicate in 96-well plates. All media used are listed in Table S3. The category 'carbon source' indicates that 2% of glucose was substituted with the indicated concentration of the specified carbon source. Likewise, the category 'nitrogen source' indicates that 0.5% of ammonia sulfate was substituted with the indicated concentration of the specified nitrogen source. Strains were cultured during 60 h and $OD_{600\ nm}$ was measured every 4 h using a microplate reader (Varioskan Flash Multimode Reader, Thermo Scientific, USA). Raw OD data were processed as follows: OD measurements of non-inoculated media were subtracted and then, non-linearity of high-density cultures were corrected using the formula $OD_{correc} = 0.2453(OD_{obs})^3 + 0.2735(OD_{obs})^2 + 0.9779(OD_{obs}) - 0.0577$ (Warringer & Blomberg, 2003). Finally, wrong data were removed to obtain smoothed growth data, making easy the growth parameters to be extracted. Fitness of the strains was analysed by extracting the growth parameters (lag time, growth rate and proliferative efficiency) from the growth curves using GrowthRates R package (Hall *et al*., 2014). Growth curves were adjusted to a Baranyi model (Baranyi and Roberts, 1994). Growth parameters, extracted from growth curves, for each strain and condition are shown in Supporting Information File S4. Lag phase determination was no possible in all the growth curves (values in red in Supporting Information File S4), because of the curve shape showed in some media. These missing data were not included in the statistical analysis, neither in the heat map representation.

Microvinifications assays. The 30 wine strains were assayed in microvinification assays to characterize the effect of *IRC7-genotype* on wine parameters. Firstly, strains were precultured during 48 h in YNB-G medium (0.17% Yeast Nitrogen Base (BD Difco™, USA) and 2% glucose) in 15 ml flasks with 10 ml of medium at 28°C with shaking (120 rpm). Then, strains were inoculated at a final cell concentration of 10^6 cells ml^{-1} in Synthetic Grape Must medium described by Henschke and Jiranek (1993), that mimics a natural grape must, supplemented with 300 mg L^{-1} of sodium metabisulfite and adjusted to pH 3.5. Assays were performed in triplicate in 30 ml flasks with 25 ml of must. Fermentations were performed at 25°C with shaking at 120 rpm. Once fermentations finished, cultures were centrifugated at 7000 rpm for 10 min to remove biomass. Then, supernatants were stored at −20°C until further analysis. To quantify basic parameters of finished fermentation we used the near infrared spectroscopy method, utilizing one monochromator instrument, Foss NIRSystems 6500 SY-I (Silver Spring, MD, USA). Ethanol content, residual sugars, density, pH, malic acid, lactic acid, volatile acidity and total acidity were determined by this method.

Competition assays. To confirm the effect of *IRC7*-genotype on the fitness of the *S. cerevisiae* strains, competitions assays were carried out where two strains, harbouring the two different version of *IRC7*, were inoculated in mixed cultures. Four HF strains (HF1, HF2, HF8 and HF9) and three HS strains (HS4, HS6 and HS9) were selected for the pairwise competition assays. These experiments were performed into three different media to simulate wine fermentation, beer fermentation and non-fermentative condition. Synthetic Grape Must (Henschke and Jiranek, 1993) supplemented with 300 mg L^{-1} of sodium metabisulfite, Beer Must (malt extract 13.2%, pH 5.2) and non-fermentative medium (glycerol 6% – as sole carbon source; YNB without amino acid and ammonium sulfate 0.017% (BD Difco™, USA); CSM 0.077% (MP Biomedicals™, UK); ammonium sulfate 0.5%) were used as culture media. Cultures were performed, in triplicate, in 40 ml flasks with 30 ml of the corresponding medium, and incubated at 25°C under orbital shaking at 120 rpm. Previously, strains were precultured during 48 h in YNB-G medium. Thereafter, strains were inoculated in the proportion 1:1, reaching a final concentration of $2{\cdot}10^6$ cells ml^{-1}. Weight loss was monitored daily to determine the end of the fermentation. Final time samples were taken, serially diluted and plated to obtain colonies after incubation at 28°C. Ten colonies were selected from each replicate, and the implantation of the strains was monitored by DNA extraction and PCR amplification of the interdelta polymorphism fingerprinting method (Legras and Karst, 2003). Thus, the implantation percentage (final prevalence of each strain) and the success rate (number of competition experiments won – implantation higher than 50% – by HS strains divided by the total number of competitions) of each strain at the end of the culture were calculated.

Killer activity assays. Killer activity and killer sensitivity of the 30 strains of the study (representing the three *IRC7-genotype* group) were measured using the method

described by Santos and colleagues (2009). Strains to be tested for killer activity were inoculated in ~1 cm diameter concentrated zones onto YMA-MB plates (1% glucose, 0.3% yeast extract, 0.3% malt extract and 0.5% proteose peptone, supplemented with 30 mg L^{-1} of methylene blue, 3% NaCl and 2% agar, pH 4.8) previously seeded with a lawn ($5.0{\cdot}10^5$ cells ml^{-1}) of the strains to be tested for killer sensitivity. Plates were incubated for a week at 20°C. After that, killer activity was detected by the observation of the halo of inhibition. Semiquantitative estimation of the killer activity intensity was calculated as the total diameter of the halo (growth inhibition) divided by the diameter of the biomass concentrated zone.

Oxidative stress evaluation. To compare the oxidative stress level in fermentation conditions between both *IRC7*^{F} and *IRC7*^{S} genotype group, intracellular ROS levels were evaluated. As described in the high throughput phenotyping, strains were precultivated and inoculated in SGM medium. After 24 h of fermentation, cells were treated during 90 min with menadione 1 mM. Then, cells were treated with dihydrorhodamine 123 (DHR 123, Sigma-Aldrich) at a final concentration of 7.5 μg ml^{-1}, and incubated in the dark during 90 min at 28°C under orbital shaking. After that, cells were pelleted, washed and resuspended in PBS. Then, the oxidative stress was analysed by quantified green florescence emission (540 nm) after excitation (485 nm) in a microplate reader Varioskan Flash Multimode Reader (Thermo Scientific).

Pseudohyphal growth test. For pseudohyphal growth development, yeasts were grown on minimal medium (0.17% yeast nitrogen base without amino acids and ammonium sulfate, 2% glucose, and 10 mM ammonium sulfate) for 16 h at 28°C. After that, cells were harvested and diluted (10^6 factor). From these dilutions, 100 μl were taken and spread onto solid Synthetic Low-Ammonium-Dextrose (SLAD) medium (0.17% yeast nitrogen base without amino acids and ammonium sulfate, 2% glucose, and 50 μM ammonium sulfate). Plates were incubated for 5 days at 28°C and colonies were observed by microscopy (10×).

Genome sequencing and identification of IRC7-associated mutations

Nine *S. cerevisiae* wine strains genomes, in representation of the three *IRC7* genotypes, were sequenced (GenBank accession PRJNA646611). Total genomic DNA was extracted using ZR Fungal/Bacterial DNA MiniPrep Kit (Zymo Research®, USA). Library preparation was performed by enzymatic fragmentation using Nextera DNA Library Prep kit (Illumina, USA). Libraries were sequenced using Illumina NextSeq 500/550 v2.5 per kits, with a coverage of 50×, obtaining sequences with the following stats; total number of reads from 6,876,948 to 11,022,748; total bases sequenced from 1,018,248,534 to 1,637,804,639; and an average length of 148 bases (minimum length of 35 bases and maximum length of 151 bases).

Sequence was filter with Trimmomatic v.0.38 (Bolger *et al*., 2014) with the following parameters: sliding window 15, minimum quality Q20 and minimum length 140. Filtered sequence was aligned with BWA (0.7.15-r1140; Li and Durbin, 2010) against *S. cerevisiae* VL3 genome (GenBank accession GCA_000190235.1_ScVL3_v01). Variant calling process was performed using GATK4 (4.0.4.0; McKenna *et al*., 2010). The process includes: indel realignment, duplicate removal, and performed SNP and INDEL discovery. The parameter was set according to GATK Best Practices recommendations. SNP and INDEL functional annotations were performed by SnpEFF4.3 t (Cingolani *et al*., 2012).

Variant calling results were treated with R version 3.6.3 (R Development Core Team, 2019) and the packages vcfR (Knaus and Grünwald, 2017) and tidyverse (Wickham *et al*., 2019). Bipartite network was built selecting the combination of chromosome, position and alternative and linking with the strain. The resulting network was visualized using the Gephi software version 0.9.2. (Bastian *et al*., 2009). Only 'moderate' and 'high importance' sequence variants identified by SnpEFF4.3t – according to its putative functional impact – have been taken into account for further analysis (see http://snpeff.sourceforge.net/VCFannotationformat_v1.0.pdf for more detailed information about this categories).

Genomic survey

In order to study the distribution of the previous identified mutations across the *S. cerevisiae* population, genomic results were combined with all genomes from the publication of Gallone and colleagues (2016) (including *S. cerevisiae* genomes from six different phylogenetic populations). In order to compare and combine the results original fastq sequences were downloaded from ENA database. All samples were processed with the same workflow as described above (except for the functional annotation). All tables were processed using R and tidyverse packages. Mutation heatmap was performed using pheatmap (Kolde, 2019). The co-occurrences with *IRC7*^{S} and *IRC7*^{F} allelic variants of each mutation were obtained using the results from mutations distribution data (Supporting Information File S2). The presence of the relevant allelic variants identified in this work were double-checked in the *S. cerevisiae* strains listed in Table S2, by the use of specific primers for the

amplification and sequencing of the regions containing these mutations (Table S6).

Statistical analysis

Statistical analysis was performed with the package stats of R software, version 3.6.3 (R Development Core Team, 2019). Analysis of variance (ANOVA) and Tukey post-hoc tests were applied to compare means of the different assays. *T*-test were applied to compare implantation percentages resulting from the (HS vs. HF) competition assays. Principal component analysis were applied to analyse wine basic parameters of microvinifications assays using the prcomp package in R.

Acknowledgements

Funding for the research in this paper was provided by Agrovin S.A, under the framework of the CDTI-financed project IDI-20160102, by Pago de Carraovejas S.L.U. CDTI-financed project IDI-20160750 (Centre for Industrial Technological Development, Spanish Ministry of Economy, Industry and Competitiveness, Spain), by the Madri+d foundation and the Rey Juan Carlos University (R&D projects for young researchers under the competitive project 'WiNetworks') and by the Ministry of Science and Innovation of Spain (MICINN project 'Wineteractions': PID2019-105834GA-I00).

Javier Ruiz acknowledges to the Complutense University of Madrid for his doctoral grant (CT17/17-CT18/17) and Miguel de Celis acknowledges to the Spanish Ministry of Economy, Industry and Competitiveness for his doctoral grant (BES-2017-080024) in the framework of the project CTM2016-76491-P.

References

Almeida, P., Barbosa, R., Bensasson, D., Gonçalves, P., and Sampaio, J.P. (2017) Adaptive divergence in wine yeasts and their wild relatives suggests a prominent role for introgressions and rapid evolution at noncoding sites. *Mol Ecol* **26**: 2167–2182. https://doi.org/10.1111/mec.14071.

Almeida, P., Barbosa, R., Zalar, P., Imanishi, Y., Shimizu, K., Turchetti, B., *et al*. (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.

Baranyi, J., and Roberts, T.A. (1994) A dynamic approach to predicting bacterial growth in food. *Int J Food Microbiol* **23**: 277–294. https://doi.org/10.1016/0168-1605(94)90157-0.

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

Bastian, M.., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

Belda, I., Gobbi, A., Ruiz, J., de Celis, M., Ortiz-Álvarez, R., Acedo, A., and Santos, A. (2020) Microbiomics to define wine terroir. In *Reference Module in Food Science*. Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-08-100596-5.22875-8.

Belda, I., Ruiz, J., Beisert, B., Navascués, E., Marquina, D., Calderón, F., *et al*. (2017) Influence of *Torulaspora delbrueckii* in varietal thiol (3-SH and 4-MSP) release in wine sequential fermentations. *Int J Food Microbiol* **257**: 183–191. https://doi.org/10.1016/j.ijfoodmicro.

Belda, I., Ruiz, J., Navascués, E., Marquina, D., and Santos, A. (2016) Improvement of aromatic thiol release through the selection of yeasts with increased β-lyase activity. *Int J Food Microbiol* **225**: 1–8. https://doi.org/10.1016/j.ijfoodmicro.2016.03.001.

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

Breslow, D.K., Cameron, D.M., Collins, S.R., Schuldiner, M., Stewart-Ornstein, J., and Newman, H.W. (2008) A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods* **5**: 711–718. https://doi.org/10.1038/nmeth.1234.

Camarasa, C., Grivet, J.P., and Dequin, S. (2003) Investigation by 13C-NMR and tricarboxylic acid (TCA) deletion mutant analysis of pathways for succinate formation in *Saccharomyces cerevisiae* during anaerobic fermentation. *Microbiology (Reading)* **149**: 2669–2678. https://doi.org/10.1099/mic.0.26007-0.

Carstens, E., Lambrechts, M.G., and Pretorius, I.S. (1998) Flocculation, pseudohyphal development and invasive growth in commercial wine yeast strains. *S Afr J Enol Vitic* **19**: 52–61.

Cavazza, A., Guzzon, R., Malacarne, M., and Larcher, R. (2013) The influence of the copper content in grape must on alcoholic fermentation kinetics and wine quality. A survey on the performance of 50 commercial active dry yeasts. *Vitis* **52**: 149–155.

Cingolani, P., Platts, A., Wang, L., Coon, M., Nguyen, T., Wang, L., *et al*. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118, iso-2; iso-3. *Fly (Austin)* **6**: 80–92. https://doi.org/10.4161/fly.19695.

Conacher, C.G., Rossouw, D., and Bauer, F.F.B. (2019) Peer pressure: evolutionary responses to biotic pressures in wine yeasts. *FEMS Yeast Res* **19**: foz072. https://doi.org/10.1093/femsyr/foz072.

Cordente, A.G., Borneman, A.R., Bartel, C., Capone, D., Solomon, M., Roach, M., and Curtin, C.D. (2019) Inactivating mutations in Irc7p are common in wine yeasts, attenuating carbon-sulfur β-lyase activity and volatile sulfur compound production. *Appl Environ Microbiol* **85**: e02684–e02618. https://doi.org/10.1128/AEM.02684-18.

Crosato, G., Nadai, C., Carlot, M., Garavaglia, J., Ziegler, D. R., Rossi, R.C., *et al*. (2020) The impact of CUP1 gene copy-number and XVI-VIII/XV-XVI translocations on copper and sulfite tolerance in vineyard *Saccharomyces*

cerevisiae strain populations. *FEMS Yeast Res* **20**: foaa028. https://doi.org/10.1093/femsyr/foaa028.

Cullen, P.J., and Sprague, G.F. (2012) The regulation of filamentous growth in yeast. *Genetics* **190**: 23–49. https://doi.org/10.1534/genetics.111.127456.

Dunn, B., Richter, C., Kvitek, D.J., Pugh, T., and Sherlock, G. (2012) Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res* **22**: 908–924. https://doi.org/10.1101/gr.130310.111.

Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., *et al*. (2016) Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* **166**: 1397–1410. https://doi.org/10.1016/j.cell.2016.08.020.

García-Ríos, E., and Guillamón, J.M. (2019) Mechanisms of yeast adaptation to wine fermentations. *Prog Mol Subcell Biol* **58**: 37–59. https://doi.org/10.1007/978-3-030-13035-0_2.

Gasmi, N., Jacques, P.E., Klimova, N., Guo, X., Ricciardi, A., Robert, F., and Turcotte, B. (2014) Dual role for yeast Ert1: activator of gluconeogenic genes and repressor of *PDC1*, a key fermentation gene. *Genetics* **3**: 114–560. https://doi.org/10.1534/genetics.114.168609.

Gerke, J., Lorenz, K., and Cohen, B. (2009) Genetic interactions between transcription factors cause natural variation in yeast. *Science* **323**: 498–501. https://doi.org/10.1126/science.1166426.

Gonçalves, M., Pontes, A., Almeida, P., Barbosa, R., Serra, M., Libkind, D., *et al*. (2016) Distinct domestication trajectories in top-fermenting beer yeasts and wine yeasts. *Curr Biol* **26**: 750–2761. https://doi.org/10.1016/j.cub.2016.08.040.

Gutiérrez, A., Beltran, G., Warringer, J., and Guillamón, J.M. (2013) Genetic basis of variations in nitrogen source utilization in four wine commercial yeast strains. *PLoS One* **8**: e67166. https://doi.org/10.1371/journal.pone.0067166.

Hall, B.G., Acar, H., Nandipati, A., and Barlow, M. (2014) Growth rates made easy. *Mol Biol Evol* **31**: 232–238. https://doi.org/10.1093/molbev/mst187.

Harsch, M.J., and Gardner, R.C. (2013) Yeast genes involved in sulfur and nitrogen metabolism affect the production of volatile thiols from Sauvignon Blanc musts. *Appl Microbiol Biotechnol* **97**: 223–235. https://doi.org/10.1007/s00253-012-4198-6.

Henschke, P.A., and Jiranek, V. (1993). Yeasts–metabolism of nitrogen compounds. In G.H. Fleet (Ed.), *Wine Microbiology and Biotechnology* (Switzerland: Harwood Academic, 1993) pp 77-164

Howell, K.S., Swiegers, J.H., Elsey, G.M., Siebert, T.E., Bartowsky, E.J., Fleet, G.H., *et al*. (2004) Variation in 4-mercapto-4-methyl-pentan-2-one release by *Saccharomyces cerevisiae* commercial wine strains. *FEMS Microbiol Let* **240**: 125–129. https://doi.org/10.1016/j.femsle.2004.09.022.

Knaus, B.J., and Grunwald, N.J. (2017) VCFR: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour* **17**: 44–53.

Kolde, K. (2019). pheatmap: Pretty Heatmaps. R package version 1.0.12.

Legras, J.L., Galeote, V., Bigey, F., Camarasa, C., Marsit, S., Nidelet, T., *et al*. (2018) Adaptation of *S. cerevisiae* to fermented food environments reveals remarkable genome plasticity and the footprints of domestication. *Mol Biol Evol* **35**: 1712–1727. https://doi.org/10.1093/molbev/msy066.

Legras, J.L., and Karst, F. (2003) Optimisation of interdelta analysis for *Saccharomyces cerevisiae* strain characterisation. *FEMS Microbiol Lett* **221**: 249–255. https://doi.org/10.1016/S0378-1097(03)00205-2.

Legras, J.L., Merdinoglu, D., Cornuet, J.M., and Karst, F. (2007) Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* **16**: 2091–2102. https://doi.org/10.1111/j.1365-294X.2007.03266.x.

Li, H., and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics* **26**: 589–595. https://doi.org/10.1093/bioinformatics/btp698.

Liti, G. (2015) The natural history of model organisms: the fascinating and secret wild life of the budding yeast *S. cerevisiae*. *Elife* **4**: e05835. https://doi.org/10.7554/eLife.05835.

Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., and James, S.A. (2009) Population genomics of domestic and wild yeasts. *Nature* **458**: 337–341. https://doi.org/10.1038/nature07743.

Lu, L., Roberts, G.G., Oszust, C., and Hudson, A.P. (2005) The YJR127C/ZMS1 gene product is involved in glycerol-based respiratory growth of the yeast *Saccharomyces cerevisiae*. *Curr Genet* **48**: 235–246. https://doi.org/10.1007/s00294-005-0023-4.

Marquina, D., Santos, A., and Peinado, J.M. (2002) Biology of killer yeasts. *Int Microbiol* **5**: 65–71. https://doi.org/10.1007/s10123-002-0066-z.

Marsit, S., and Dequin, S. (2015) Diversity and adaptive evolution of *Saccharomyces* wine yeast: a review. *FEMS Yeast Res* **15**: fov067. https://doi.org/10.1093/femsyr/fov067.

McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R., Moreau, R.A., *et al*. (2004) Fermented beverages of pre- and proto-historic China. *PNAS* **101**: 17593–17598. https://doi.org/10.1073/pnas.0407921102.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., *et al*. (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**: 1297–1303. https://doi.org/10.1101/gr.107524.110.

Peter, J., De Chiara, M., Friedrich, A., Yue, J.X., Pflieger, D., Bergström, A., *et al*. (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., and Sampaio, J.S. (2020) Revisiting the taxonomic synonyms and populations of *Saccharomyces cerevisiae*-phylogeny, phenotypes, ecology and domestication. *Microorganisms* **8**: 903. https://doi.org/10.3390/microorganisms8060903.

Rivero, D., Berná, L., Stefanini, I., Baruffini, E., Bergerat, A., Csikász-Nagy, A., *et al*. (2015) Hsp12p and *PAU* genes are involved in ecological interactions between natural yeast strains. *Environ Microbiol* **17**: 3069–3081. https://doi.org/10.1111/1462-2920.12950.

Roncoroni, M., Santiago, M., Hooks, D.O., Moroney, S., Harsch, M.J., Lee, S.A., Richards K.D., Nicolau L., Gardner R.C. (2011). The yeast *IRC7* gene encodes a β-lyase responsible for production of the varietal thiol 4-mercapto-4-methylpentan-2-one in wine. *Food Microbiol* **28**:926– 935 https://doi.org/10.1016/j.fm.2011.01.002, 28, 5, 926, 935

Santiago, M., and Gardner, R.C. (2015) The *IRC7* gene encodes cysteine desulphydrase activity and confers on yeast the ability to grow on cysteine as a nitrogen source. *Yeast* **32**: 519–532. https://doi.org/10.1002/yea.3076.

Santos, A., San Mauro, M., Bravo, E., and Marquina, D. (2009) PMKT2, a new killer toxin from *Pichia membranifaciens*, and its promising biotechnological properties for control of the spoilage yeast *Brettanomyces bruxellensis*. *Microbiology (Reading)* **155**: 624–634. https://doi.org/10.1099/mic.0.023663-0.

Schacherer, J., Shapiro, J.A., Ruderfer, D.M., and Kruglyak, L. (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* **458**: 342–345. https://doi.org/10.1038/nature07670.

Schmidt, S.A., Kolouchova, R., Forgan, A.H., Borneman, A. R. (2020) Evaluation of *Saccharomyces cerevisiae* wine yeast competitive fitness in enologically relevant environments by barcode sequencing. *G3* **10**: 591–603. https://doi.org/10.1534/g3.119.400743, 10, 2, 591, 603

Servien□, E., Lukša, J., Orentait□, I., Lafontaine, D.L., and Urbonavičius, J. (2012) Screening the budding yeast genome reveals unique factors affecting K2 toxin susceptibility. *PLoS One* **7**: e50779. https://doi.org/10.1371/journal.pone.0050779.

Steensels, J., Gallone, B., Voordeckers, K., and Verstrepen, K.J. (2019) Domestication of industrial microbes. *Curr Biol* **29**: R381–R393. https://doi.org/10.1016/j.cub.2019.04.025.

Strope, P.K., Skelly, D.A., Kozmin, S.G., Mahadevan, G., Stone, E.A., Magwene, P.M., *et al*. (2015) The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res* **25**: 762–774. https://doi.org/10.1101/gr.185538.114.

Swiegers, J., and Pretorius, I. (2007) Modulation of volatile sulfur compounds by wine yeast. *Appl Microbiol Biotechnol* **74**: 954–960. https://doi.org/10.1007/s00253-006-0828-1.

Tominaga, T., Furrer, A., Henry, R., and Dubourdieu, D. (1998) Identification of new volatile thiols in the aroma of *Vitis vinifera L*. var. Sauvignon blanc wines. *Flav Frag J* **13**: 159–162.

Vermeersch, L., Perez-Samper, G., Cerulus, B., Jariani, A., Gallone, B., Voordeckers, K., *et al*. (2019) On the duration of the microbial lag phase. *Curr Genet* **65**: 721–727. https://doi.org/10.1007/s00294-019-00938-2.

Walker, M.E., Nguyen, T.D., Liccioli, T., Schmid, F., Kalatzis, N., Sundstrom, J.F., *et al*. (2014) Genome-wide identification of the Fermentome; genes required for successful and timely completion of wine-like fermentation by *Saccharomyces cerevisiae*. *BMC Genomics* **15**: 552. https://doi.org/10.1186/1471-2164-15-552.

Wang, Q.M., Liu, W.Q., Liti, G., Wang, S.A., and 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.

Warringer, J., and Blomberg, A. (2003). Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in Saccharomyces cerevisiae. *Yeast*, **20**: 53–67.

Warringer, J., Zörgö, E., Cubillos, F.A., Zia, A., Gjuvsland, A., Simpson, J.T., *et al*. (2011) Trait variation in yeast is defined by population history. *PLoS Genet* **7**: e1002111. https://doi.org/10.1371/journal.pgen.1002111.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., *et al*. (2019) Welcome to the tidyverse. *J Open Source Softw* **4**: 1686. https://doi.org/10.21105/joss.01686.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supplementary Information

Table S1. Mutations shared amongst the HS strains (highlighted in red in Fig. S1) identified in the variant calling analysis of nine of our studied strains (including three representative strains of each IRC7-genotype (HS4. HS6. HS9; HT3. HT6. HT10; HF1. HF2. HF9) and VL3 as a reference strain.

Table S2. *S. cerevisiae* strains used in the phenotyping study.

Table S3. Culture media used in the *S. cerevisiae* strains phenotyping.

Table S4. Enological parameters at the end of the microvinification assays.

Table S5. Average lag time, growth rate and proliferative efficiency of the yeast strains pertaining to different genotype groups (homozygous for the reference allele (HR), heterozygous (HT) and homozygous for the alternative variant allele (HV)), according to allelic variants found in different genes highlighted in Fig. 2B, assayed in a panel of media (Table S3). Different letters indicate the existence of significant differences in an ANOVA comparison. Comparisons with significant differences are highlighted with bold letters.

Table S6. List of primers used to confirm some relevant allelic variants found in this work.