Yeast
Yeast 2016; **33**: 277–287
Published online in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/yea.3173

ISSY32 Special Issue

Social wasp intestines host the local phenotypic variability of *Saccharomyces cerevisiae* strains

Leonardo Dapporto[1†], Irene Stefanini[2†], Damariz Rivero[3], Mario Polsinelli[4], Paolo Capretti[5], Paolo De Marchi[6], Roberto Viola[2], Stefano Turillazzi[4,7] and Duccio Cavalieri[2,4*]

[1]*Instituto de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Spain*
[2]*Department of Computational Biology, Centre for Research and Innovation, Fondazione Edmund Mach, San Michele all'Adige, Trento, Italy*
[3]*NEUROFARBA Department, University of Florence, Italy*
[4]*Department of Biology, University of Florence, Italy*
[5]*DISPAA Department, University of Florence, Italy*
[6]*Cantina Isole e Olena, Florence, Italy*
[7]*Centro di Servizi di Spettromeria di Massa, University of Florence, Italy*

**Correspondence to:*
D. Cavalieri, Deparment of Biology University of Florence Via Madonna del Piano 6-50019 Sesto Fiorentino, Florence, Italy.
E-mail: duccio.cavalieri@unifi.it

†These authors equally contributed to this study.

Received: 30 November 2015
Accepted: 4 May 2016

Abstract

Nowadays, the presence of *Saccharomyces cerevisiae* has been assessed in both wild and human-related environments. Social wasps have been shown to maintain and vector *S. cerevisiae* among different environments. The availability of strains isolated from wasp intestines represents a striking opportunity to assess whether the strains found in wasp intestines are characterized by peculiar traits. We analysed strains isolated from the intestines of social wasps and compared them with strains isolated from other sources, all collected in a restricted geographic area. We evaluated the production of volatile metabolites during grape must fermentation, the resistance to different stresses and the ability to exploit various carbon sources. Wasp strains, in addition to representing a wide range of *S. cerevisiae* genotypes, also represent large part of the phenotypes characterizing the sympatric set of yeast strains; their higher production of acetic acid and ethyl acetate could reflect improved ability to attract insects. Our findings suggest that the relationship between yeasts and wasps should be preserved, to safeguard not only the natural variance of this microorganism but also the interests of wine-makers, who could take advantage from the exploitation of their phenotypic variability. Copyright © 2016 John Wiley & Sons, Ltd.

Keywords: **phenotype; ecology; saccharomyces cerevisiae; social wasps**

Introduction

Saccharomyces cerevisiae is one of the most-used microorganisms, exploited by humans over the last 9000 years in food and beverage production (McGovern *et al*., 2004). It is surprising that, despite the wide knowledge obtained by the use of this microorganism in several biological topics, understanding of its genetic and phenotypic potentials is only in its early stages. Such a lack of knowledge is mainly due to the late recognition of *S. cerevisiae* as a 'wild' organism (Liti *et al*., 2009), residing in habitats different from those associated with human activities. *S. cerevisiae* is known for its ability to ferment grape must and confer to the fermentation product peculiar aromatic characteristics. Strains have been largely selected and used as inocula to conduct alcoholic fermentation and prevail over other natural microbiota present in fresh must. *S. cerevisiae* strains characterized by particular abilities to grow in must and to produce pleasant metabolites have been selected by wine-makers since the nineteenth century (Muller-Thurgau, 1896). Hyma *et al*. (2011) suggested that the

Yeast

differences in fermentation aromas produced by wine and wild strains can be perceived by humans, thus providing a fundamental contribution to explaining the effect of human selection on yeast phenotypes. The same authors also concluded that every individual descending from the wine ancestor is probably a 'pleasant fermentor'. The authors' assumption of the phenotypic uniformity of the wine clade was justified by the fact that, at the time of their study, almost every strain of this clade was isolated from wineries, and most of them were artificially selected because of their ability to carry out pleasant fermentations. Recently, this genetic cluster has been broadened after the isolation of strains belonging to different sources and geographic regions (Fay and Benavides, 2005; Legras *et al*., 2007; Liti, 2015; Liti *et al*., 2009). Comparison of strains isolated worldwide represents an unprecedented possibility to describe *S. cerevisiae* populations, but it has to be carefully considered that the variation over geographic locations could play a relevant role in defining yeast evolution. It cannot be excluded that at smaller geographic scales sympatric wild and human-related strains show a different pattern of phenotypic divergence, due to convergent local adaptations and/or gene exchanges between wild and wine strains. Knight *et al*. (2015) made a step forward in distinguishing niches by geographic location. Indeed, they showed that *Saccharomyces cerevisiae* strains isolated from different sources around vineyards at a scale of >100 km in New Zealand can produce a vineyard-specific aromatic bouquet of ferments. The same strains also bore a genotype differing according to the geographic origin of the strains, rather than to their isolation source (Knight and Goddard, 2015). These results suggest that the fermentation phenotypes expressed by strain are more related to their genetics than to the hypothetical selective pressures they experience in different sources of isolation.

S. cerevisiae strains have been shown to be vectored among different environments by animals (Francesca et al., 2012; Goddard *et al*., 2010; Hyma and Fay, 2013; Wang *et al*., 2012). Such movements may have mixed strains bearing different phenotypes, thus preventing the identification of traits specific for each isolation source. Our group has recently shown that *S. cerevisiae* can be maintained in natural environments and vectored by social wasps (Stefanini *et al*., 2012). The same study revealed that yeasts can persist inside wasp guts for a long time, due to trans-generational transmission among adult wasps and their offspring. Yeasts found in wasp guts bear all the local yeast genetic variance, descending both from wild and wine ancestors.

The association of *S. cerevisiae* and social wasps and the observation that wasps host yeasts all year long represent a striking opportunity to dissect the phenotypic variability of this microorganism. Since social wasps are vectors of this yeast, whichever is the direction of the spread of microorganisms (either from the cellar to the vineyard or vice versa), the yeast strains present in wasp intestines could represent all the phenotypes observed in a given geographical area. We have thus established an experimental design taking advantage of two main points: (a) the use of sympatric strains allowing the removal of a confounding systematic bias associated with a population structure produced by geographical distances; (b) the availability of strains directly isolated from the intestines of social wasps (without culture enrichment), allowing exploration of the phenotypic variability of strains found in natural vectors.

The main aim of the study was to ascertain whether strains occurring in the vector show specific traits. With this in mind, we analysed several phenotypes, ranging from the production of volatile metabolites to resistance to stress and the ability to exploit carbon sources in a set of *S. cerevisiae* strains isolated from different sources of a restricted geographic area. This area encompasses the Chianti hills (Tuscany, Italy), an important centre of origin for *S. cerevisiae* human-related strains, where wine fermentation is documented to have been in use since at least the third century BC (Robinson, 2006). We also included strains isolated from the intestine of wasps collected on Elba Island (Tuscan archipelago), the vast majority of whose territory was dedicated to vineyards until the recent past (40–50 years ago). Strains have been compared in phenotypes with strains widely used for laboratory and wine fermentation purposes. The availability of strains isolated from wasp guts belonging to the wine cluster and to other genetic clades allowed measurement of the relevance of ancestry to the fermentation characteristics of these strains.

DOI: 10.1002/yea

Materials and methods

S. cerevisiae strains

Ten *S. cerevisiae* strains isolated from social wasp intestines were selected as representative of the genetic variability among the strains isolated from this source (Stefanini *et al*., 2012). In addition, 17 strains isolated from grape skin and wine fermentation were used as reference for the characterization of the phenotypes (Table 1). All 27 strains (except the laboratory strain) were collected in a restricted geographical area of Tuscany (Chianti and Elba Island). The S288c laboratory strain, commonly used as reference, was excluded, since it has been shown to be significantly different in several traits in comparison to natural yeast strains (Warringer *et al*., 2011). Three genes, *IRC8*, *EXO5* and *URN1*, were sequenced and analysed as previously described by Ramazzotti *et al*. (2012); these genes were chosen as able to recapitulate the whole genome phylogeny (Ramazzotti *et al*., 2012). Population ancestries were estimated using the model-based program STRUCTURE (Pritchard *et al*., 2000). The results of 10 independent STRUCTURE chains were combined with CLUMPP (Jakobsson and Rosenberg, 2007). The identified population structure was used to classify strains according to their ancestries into the wine/European clean population (CP) and rare genomes (RG).

Characterization of fermentation ability of strains

The investigated strains belong to vineyards where grapes are used to produce both red (mainland) and white (Elba Island) wines. To fully characterize the fermentation characteristics of these strains, we evaluated growth abilities and metabolite profiles in filtered and sterilized red and white grape musts. After a preculture

Table 1. Strains origin, genotype and isolation source

Strain	Origin	Genomic cluster	Isolation type
BIBVC1_1	Isole e Olena (Tuscany, Italy)	RG	Wasp
BIBVC4_3	Isole e Olena (Tuscany, Italy)	RG	Wasp
BIBVC5_3	Isole e Olena (Tuscany, Italy)	RG	Wasp
Buc1	Mercatale Val di Pesa (Florence, Italy)	CP	Wasp
CIU8	Chianti (Florence, Italy)	RG	Wasp
CPT2	Mercatale Val di Pesa (Florence, Italy)	RG	Wasp
E32	Mercatale Val di Pesa (Florence, Italy)	CP	Wasp
E4	Mercatale Val di Pesa (Florence, Italy)	RG	Wasp
F31x	Tignano (Florence, Italy)	RG	Wasp
NPSM	Pelago (Florence, Italy)	RG	Wasp
Reg1b	Monsanto, Barberino Val d'Elsa (Florence, Italy)	RG	Wasp
Sgu428R	Chianti (Florence, Italy)	CP	Grape
Sgu52	Chianti (Florence, Italy)	CP	Grape
Sgv114	Chianti (Florence, Italy)	RG	Wasp
starter	Isole e Olena (Tuscany, Italy)	CP	Wine
VAT11	Isole e Olena (Tuscany, Italy)	CP	Wine
VAT13	Isole e Olena (Tuscany, Italy)	CP	Wine
VAT16	Isole e Olena (Tuscany, Italy)	CP	Wine
VS138	Chianti (Florence, Italy)	RG	Grape
VS180	Chianti (Florence, Italy)	RG	Grape
VS274	Chianti (Florence, Italy)	CP	Grape
VS290	Chianti (Florence, Italy)	RG	Grape
W303	Established by R. Rothstein	RG	Laboratory
YVC1E2	Isola d'Elba (Tuscany, Italy)	CP	Wasp
YVC2E6	Isola d'Elba (Tuscany, Italy)	CP	Wasp
YVC4EST1	Isola d'Elba (Tuscany, Italy)	CP	Wasp
YVPC7.6	Grassina (Florence, Italy)	RG	Wasp

RG, rare genomes; CP, wine/European clean population.

DOI: 10.1002/yea

in yeast peptone dextrose (YPD) (Radford, 1991), yeast cells were inoculated in 5 ml must (10^7 cells/ml) and maintained at 27 °C in static culture. Growth ability of the strains in musts was evaluated as amount of produced CO_2 (g/100 ml mass loss), daily measured during the fermentation. At least three independent replicates were carried out for each strain. Fermentation was halted by removing yeast cells by centrifugation.

The production of the well-known volatile metabolites that modify wine aroma, ethyl acetate, 2-methyl-1-propanol, 3-methyl-1-butanol, 2-methyl-1-butanol, 3-hydroxy-2-butanone (acetoin), phenyl ethyl alcohol, acetaldehyde, 1-propanol, acetic acid and methyl alcohol (Romano *et al.*, 1995, 1998; Styger *et al.*, 2011; Wondra and Berovic, 2001), was quantified. The analyses were performed using a GC–MS instrument (Varian, Agilent Technologies, Milan, Italy) consisting of a CP3800 gas chromatograph coupled with a 4000 ion trap MS, equipped with a SPB-624 (Supelco, Milan, Italy) capillary column (30 m × 0.25 mm i.d., 1.4 µm film thickness). The fermented musts were analysed by headspace–SPME with a 75 µm Carboxen/PDMS fibre (Supelco). The compound extraction was performed from a sealed 2 ml vial containing 0.1 ml must; the adsorption time was 3 min at room temperature and the desorption time was 10 min in the injection port at 250 °C. Helium at 1.0 ml/min was used as the carrier gas and injections were performed in splitless mode (2 min). The oven was maintained at 40 °C for 2 min, then the temperature was increased to 64 °C at 1.5 °C/min, then to 180 °C at 12 °C/min, and then maintained for 10 min (total run time 29 min). The 4000 ion trap operated in electron ionization mode with an external source; MS temperatures set for acquisition were: 100 °Ctrap, 50° manifold, 170 °C transfer line and 180 °C ion source. All the peaks were identified from their mass spectra by comparison with NIST libraries (v. 2.0f) and with chemical standards. Commercial standards of each compound, dissolved in must, were used to confirm the identification and to calibrate the GC–MS data to obtain quantitative measurements. Methyl alcohol-d3, phenylethanol-d5 and acetic acid-d3 were used as internal standards.

Strains growth abilities and resistance to stresses

The ability to metabolize various carbon sources was assessed by plating cells (10^2, 10^3, 10^4 and 10^5 cells) onto solid YP medium (1% yeast extract, 2% peptone, 2% agar) with added 2% glucose, galactose, ethanol or lactose (Hampsey 1997). Then cell growth was scored after 5 days as 0, 0.5 and 1 (non-growing, slow-growing and growing, respectively). The growth in minimal medium was assessed on solid yeast nitrogen base (YNB), eventually with added ammonium sulphate and/or amino acids. Resistance to stresses (copper, oxidative stress and ethanol) was assessed by evaluating the minimal inhibitory concentration (MIC) in YPD solid medium with several added concentrations of $CuSO_4$ (1, 2.5, 5 and 10 mm), H_2O_2 (5, 10, 15, 20, 25 and 30 mm) and ethanol (12%, 14%, 16%, 18%, 20% and 22%). The sporulation rate after 4 days in YPD (YP with added 2% dextrose) preculture was assessed as previously described (Radford, 1991). The ability of yeast cells to invade agar (Vopalenska *et al.*, 2005) and to grow at different temperatures (6 °C, 27 °C, 37 °C and 40 °C) were also assessed. The ability of strains to form hyphae was microscopically assessed after 3 days of culture in either YPD or YNB at 28 °C or 37 °C.

Statistical analyses

Differences among strains grouped by isolation source were inspected for each trait separately. Measured phenotypes were corrected for the strains' genetic backgrounds, as previously descibed (Yu et al., 2006). Volatile compounds measured in red and white must fermentations were standardized before comparison (zero mean, unit variance). To identify chemicals that could account for discrimination among strains, we then performed a partial least squares discriminant analysis (PLSDA) using the mixOmics R package (Le Cao *et al.*, 2009; Borcard *et al.*, 2011), both corrected and not corrected for population structure. To verify whether interstrain dissimilarities were larger than the intrastrain ones, we computed a dissimilarity matrix on chemical data using the Bray–Curtis index; then we compared intra- with interstrain similarities. Principal coordinate analysis (PCoA) was carried out on the measured growth ability and stress response phenotypes separately, both corrected and not

DOI: 10.1002/yea

corrected for population structure. Factor analysis on mass loss data over the measured fermentation days and obtained factors with eigenvalue > 1 was carried out to use the resulting components as variables to be included in the set of growth abilities for the following analyses. To compare strains by taking into account all the measured phenotypes, we applied PCoA on all the measured traits, both corrected and not corrected for population structure. We combined standardized measures (zero mean, unit variance) for different phenotypic classes.

To verify correspondence between phenotypes and genetic relationships, we constructed neighbour-joining trees, based on distances computed with the Kimura two parameters metric, on the concatenated SNP sequences of the selected genes. For each class of phenotype (metabolites, stresses, growth abilities) we computed a Bray–Curtis dissimilarity matrix, using all the strains and the virtual cases representing nodes in the neighbour-joining tree. We inferred the values for nodes as the mean values of the variables between the two descendants linked by a node (cases or nodes). We performed NMMDS using the ecodist package (Goslee and Urban, 2007) on the dissimilarity matrices, then phylogeny was drawn on the map obtained. Long and highly crossing branches indicate weak correlations between genetics and phenotypes. For metabolic profiles, we computed median composition among the ferment replicates for each strain. We also searched for univariate correlations among single genetic and phenotypic dissimilarity matrices by applying the Mantel test with 9999 iterations.

Wilcoxon rank sum test followed by multiple test false discovery rate (fdr) correction was carried out to compare the traits among groups of strains. When comparing strains grouped according to their isolation source, the Wilcoxon test was carried out on phenotypes (either single phenotypes or multivariate coordinates) corrected for the strains' population structures. When comparing strains grouped according to their genotypes, the Wilcoxon test was carried out on non-corrected phenotypes. Similarly, we applied the Levene test, followed by fdr correction, to compare the variance among groups of strains.

Results

According to STRUCTURE analysis on the SNP sequences of the three genes reproducing the whole-genome sequences (Ramazzotti *et al.*, 2012), *S. cerevisiae* strains isolated from various sources encompassed both isolate descendants of the locally most represented ancestor (CP) and isolates belonging to different ancestors or showing mosaic genomes (RG) (Figure 1, Table 1; see also supporting information, Figure S1).

Must fermentation – volatile compounds

The gas chromatography/mass spectrometry (GC–MS) analysis revealed the presence of nine of the 10 reported volatiles, highly matching the well assessed spectrum of *S. cerevisiae* fermentation metabolites, even though we did not find the previously recorded acetoin (3-hydroxybutanone) (Callejon *et al.*, 2010; Romano *et al.*, 1995). Among the quantified compounds, acetic acid was produced in significantly higher amounts by wasp strains than by wine strains fermenting red musts (Wilcoxon rank sum test; see supporting information, Figure S2). Similarly, the amount of ethyl acetate produced during white must fermentation was higher for wasp strains than that produced by wine strains. These differences were confirmed even after phenotype correction for population structure (see supporting information, Figure S2). The correction for population structure also allowed the detection of a higher production of ethyl acetate and methanol by grape strains compared to wasp strains during fermentations of red and white musts, respectively (see supporting information, Figure S2). PLSDA revealed that most of the measured metabolites were involved in the first component (Figure 1b). Acetic acid showed a different trend and was responsible for most of the second component (y axis; Figure 1b). 3-Methyl-1-butanol showed weights around –0.5 in the first component and around 0.5 in the second; 1-propanol and 2-methyl,1-propanol showed weights around –1 in the first component and around –0.5 in the second. Neither the first nor the second PLSDA components divided the samples according to the type of must fermented (red or white; Wilcoxon rank test, $p > 0.05$). The similarity among ferment replicates of the same

DOI: 10.1002/yea

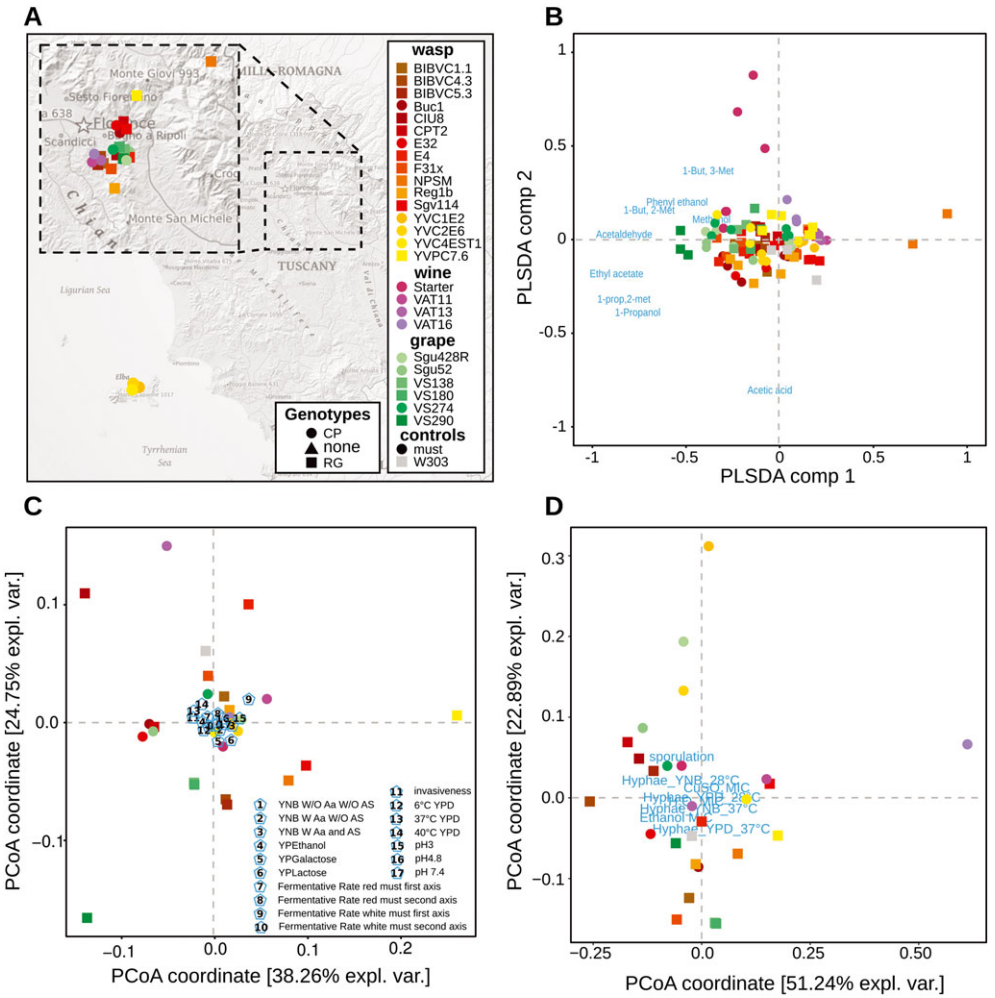

Figure 1. *Saccharomyces cerevisiae* strains origins and phenotypes. (A) *S. cerevisiae* strains classified by isolation source and genotype; CP, clear population (strains deriving from the wine ancestor); RG, rare genotype (strains deriving from different ancestors or bearing a mosaic genome). (B) First two coordinates of the PLSDA carried out on volatile compounds produced by yeasts during red and white musts fermentation, corrected for population structure. (C) First two coordinates of the PCoA carried out on yeast growth abilities, corrected for population structure. (D) First two coordinates of the PCoA carried out on yeasts response to stresses, corrected for population structure

strain was supported by the highly significant differences among intra- and inter- Bray–Curtis distances (the first among strains, the latter between ferment replicates of the same strain; $p = 2.856 \times 10^{-16}$, Wilcoxon rank sum test). Strains isolated from wasp intestines spread all along the first PLSDA component, indicating that their aromatic profiles are characterized by a wide variability in the expression of all the metabolites, with the exception of acetic acid, which was strictly linked to the second component (Figure 1b). Conversely, the second component dispersed only one of the wine strains (Starter), which showed a high value in this component. The PLSDA carried out on volatile compounds not corrected for population structure showed a similar distribution of variables, but wine strains were also spread in the second component (see supporting information, Figure S3). When compared to grape strains, wasp strains showed significantly higher values in the first PLSDA component, negatively correlated with the production of several of the volatile compounds (Wilcoxon rank sum test, $p < 0.01$; see supporting information, Figure S4, Table S1). In

DOI: 10.1002/yea

addition, wasp strains, compared to wine strains, showed significantly lower values in the second PLSDA component, negatively correlated with acetic acid production (Wilcoxon rank sum test, $p < 0.01$; see supporting information, Figure S4). The comparison of genetic groups (CP and RG) showed significant differences in the first PLSDA component (Wilcoxon rank sum test test, $p < 0.05$; see supporting information, Figure S4), with the strains descending from the wine ancestor (CP) producing higher levels of metabolites. On the other hand, the CP group showed a significantly higher variance in the second PLSDA component (Levene test $p < 0.01$; see supporting information, Figure S4), indicating that strains belonging to the wine cluster produce various levels of acetic acid. NMMDS on the metabolic profiles confirmed the overall pattern obtained with PLSDA (see supporting information, Figure S5a). The NMMDS representation slightly distorted the original dissimilarity matrix, as indicated by a moderate level of stress (s-stress = 0.131). The length and disorganization of the phylogenetic tree projected on the NMMDS coordinates showed the absence of correlation between the composition of metabolites and their phylogeny. These results were confirmed by Mantel tests indicating no significant relationships between metabolite expression and genetic distances ($p = 0.063$; see supporting information, Figure S6).

Strains' growth abilities and resistance to stresses

The factor analysis on mass loss data over the measured fermentation days resulted in two factors with eigenvalue > 1 (see supporting information, Figure S7), which were included in the set of growth ability traits. Wasp strains were found to be more sensitive to low temperatures (6 °C) than grape strains when analysing both raw data and data corrected for the strains' population structures (Wilcoxon rank sum test; see supporting information, Figure S8). The correction for population structure revealed that wasp strains show higher variance compared to wine strains in their ability to grow in the absence of ammonium sulphate (Levene test, $p < 0.05$; see supporting information, Figure S8). In addition, after the correction, wasp strains showed lower variance in their ability to grow at 6 °C compared to grape strains (Levene test, $p < 0.05$; see supporting information, Figure S8). The PCoA on growth abilities showed that the first factor relative to strains' fermentation rates in white musts was positively correlated with the first two components of the PCoA (38.26% and 24.75% explained variance) (Figure 1c). The other traits did not show strong correlations with the PCoA coordinates, except the strains' abilities to exploit galactose and lactose, which showed positive correlations with the first coordinate and negative correlations with the second coordinate (Figure 1c). The PCoA carried out on raw growth abilities (not corrected for population structure) showed a different distribution of variables (see supporting information, Figure S9), with the ferment factors showing opposite trends, according to the type of must fermented. Comparisons among genotypes (CP and RG) or isolation sources showed no differences in the PCoA components (see supporting information, Figure S9). As observed for metabolite profiles, the projection of the phylogenetic tree on the NMMDS coordinates made evident the lack of correlation between this phenotype and the individual genetics (see supporting information, Figure S5).

Considering the stress resistance traits separately, wasp strains showed a lower resistance to ethanol compared to grape strains, but this difference was not confirmed after correction of the phenotypes for population structure (see supporting information, Figure S10). The first PCoA coordinate on resistance variables was only slightly correlated with some of the traits. Contrarily, the second coordinate was positively correlated to sporulation and hyphae formation in minimal medium (YNB) at 28 °C and anti-correlated to the ability to form hyphae in rich medium (YPD) at 37 °C and to the resistance to ethanol (Figure 1d). The PCoA carried out on non-corrected phenotypes (see supporting information, Figure S11a) resulted in a different distribution of strains and variables according to the first two components, which showed a higher amount of the explained variance than the PCoA carried out on corrected phenotypes (see supporting information, Figure S1d). The Wilcoxon rank sum test did not show significant differences between strains grouped according to their genotype or isolation source (Figure S11b). As for the other groups of phenotypes, no significant correlations were observed between genetic and phenotypic distances (Mantel test; see

DOI: 10.1002/yea

supporting information, Figure S6), or a clean representation of the genetic tree on the NMDS coordinates built on stress response measurements (see supporting information, Figure S5c).

Overall phenotypic patterns

After standardizing the measures (zero mean, unit variance), we compared all the phenotypes by combining all the measured traits in a single PCoA (Figure 2a). The first two coordinates of the PCoA based on these indicators explained 31.68% and 27.47% of the total variance, respectively. Several variables were negatively correlated to the first PCoA coordinate, but the highest contribution was provided by the production of phenyl ethanol (trait 9 in Figure 2a). A few variables contributed to the second coordinate of the overall PCoA, which was negatively correlated mainly with the two axes of the fermentation factor in red musts (traits 24 and 25 in Figure 2a). As observed for other traits, the PCoA on non-corrected phenotypes resulted in a different distribution of strains and variables (see supporting information, Figure S12). Wasp strains did not show significantly different distribution or variance in the first two PCoA coordinates compared to strains isolated from other sources (Wilcoxon and Levene tests; see supporting information, Figure S12). Strains descending from the wine ancestor (CP) showed higher variance than RG strains in the second PCoA coordinate, mainly described by stress response variables (Levene test, $p < 0.01$; see supporting information, Figure S12). As observed when comparing the phenotypic classes separately, the projection of the phylogenetic tree on the NMMDS coordinates did not show any relation between individual phenotype and genetics (see supporting information, Figure S5d).

Spearman correlation analysis among phenotypes corrected for population structure was carried out for strains grouped by isolation source

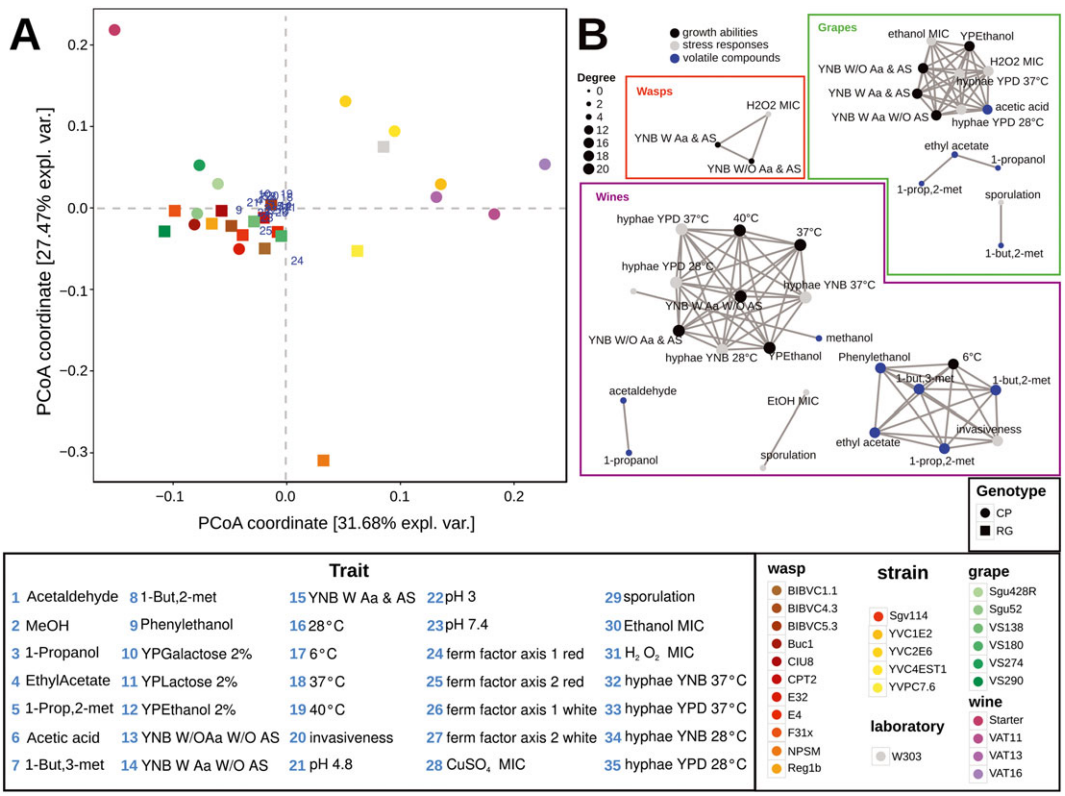

Figure 2. *Saccharomyces cerevisiae* overall phenotypic characteristics. (A) First two coordinates of the PCoA carried out on the whole set of measured phenotypes, corrected for population structure. (B) Network based on significant Spearman correlations ($r > 0.7$, $\rho < 0.05$) found among traits in strains grouped by isolation source

DOI: 10.1002/yea

and highlighted noticeable correlations among traits (Figure 2b). Wasp strains showed the lowest number of significant correlations (Spearman $\rho < 0.05$), compared to wine and grape strains. The resistance to oxidative stress was correlated with the ability to grow in minimal medium and in the absence of both amino acids and ammonium sulphate, and these traits were in turn correlated with each other. It is noteworthy that wasp strains did not show correlations among volatile compounds, which were observed in grape and wine strains. In particular, wine strains showed strong Spearman correlations among several volatile compounds and the ability to grow at low temperatures (Figure 2b). On the other hand, grape strains revealed significant Spearman correlations between several volatile compounds and among acetic acid and the strains' abilities to exploit minimal medium, ethanol as the carbon source and several responses to stress (resistance to ethanol, oxidative stress and the formation of hyphae in rich medium at both 28 °C and 37 °C) (Figure 2b).

Discussion

Genomic analyses on strains isolated all around the world revealed the existence of few clades, and strains related to human-driven fermentation fell into clear clades (Legras *et al*., 2005; Liti *et al*., 2009). At the worldwide level, the connection between phenotype and genotype of *S. cerevisiae* strains mirrors the situation observed for *Canis lupus*, one of the most-studied domesticated animals. Dogs bear higher morphological and behavioural variability with respect to that observed in the wild wolf cousin, without a corresponding increase in genetic variability (Wayne and Ostrander, 2007). Similarly, the phenotypic variability of *S. cerevisiae* is supposed to exceed that of the more genetically variable wild cousin *S. paradoxus* (Warringer *et al*., 2011). Nevertheless, these previous findings may have revealed a geographic variation of strains produced by spatial distance, rather than a real phenotypic diversity associated with the source of isolation. The analyses on our set of sympatric strains excluded the effects of the geographic distances on phenotypic variability, which could be amplified as a result of potential genetic drift and response to highly different environments (Warringer *et al*., 2011). In this small geographic-scale analysis, strains from different origins did not show clear phenotypes strictly related to the isolation source.

The pool of *S. cerevisiae* strains studied here represents the genetic variance of strains isolated in the study area, encompassing both strains descending from the wine/European ancestor and strains having a mosaic genome or descending from other ancestors. In addition, it also represents the local source heterogeneity, including strains isolated from wasp intestines, grapes and ferments.

Given the wide interest in the use of *S. cerevisiae* for wine-making purposes, the ability to ferment grape must and the organoleptic characteristics of the final product are among the most relevant and studied phenotypes of natural *S. cerevisiae* strains. As previously observed, strains belonging to the wine/European ancestor showed higher levels of volatile compounds compared to strains descending from other ancestors (Hyma *et al*., 2011). Wasp strain metabolic profiles, compared to wine strains, are characterized by a higher production of acetic acid, mainly in red must ferments. Beyond this, the metabolic profile of wasp strains does not differ from that of wine strains, indicating that wasp strains have potentials similar to those of strains selected for wine-making purposes. Indeed, when focusing on the profiles produced in white must fermentations, wasp strains produce higher levels of ethyl acetate and lower levels of methanol compared to wine strains. From the wine-making perspective, while the selected strains are obviously still the best option, since they produce lower levels of acetic acid, wasp strains could hold great potential for white must fermentations especially, even compared to grape strains, whose high production of methanol could severely impact the final product. From a biological point of view, the differential characteristics of fermentation products could explain why wasps are particularly attracted by strains with a high production of acetic acid (which is commonly used also in insect traps) and low production of methanol (a highly toxic compound) (Christiaens *et al*., 2014). This, of course, should be the object of specific future experiments.

No significant differences were found in growth abilities when considering the strain genotypes. On the contrary, wasp strains showed lower abilities to grow at low temperatures compared to grape strains, but similar to those of wine strains. In

DOI: 10.1002/yea

addition, wasp strains showed a wider range of growth in the presence of neutral pH (7.4) compared to wine strains. While the tested wine strains showed near-zero growth at this pH, some wasp strains were unable to grow under these conditions and other wasp strains showed the highest growth among all the tested strains. Considering these features, wasp and grape strains hold the potential to grow under several conditions, in the presence of different carbon sources and at various ranges of temperature and pH.

Because of their widespread residence in natural or fermentative environments, *S. cerevisiae* strains are expected to respond, react or adapt to different kind of stresses. On the contrary, strains grouped either by their isolation source or by their ancestry did not differ for any stress strait. This result could indicate two alternative scenarios: first, the isolation sources are not characterized by one or more of the tested stresses, thus the strains, even if residing there for prolonged periods, are not subjected to selection; second, strains resistant to stresses present in a specific specimen are able to escape the selection thanks as being vectored to other environments. Nevertheless, since some traits are known to characterize specific environments (i.e. high concentrations of ethanol are found in ferments only), the second hypothesis seems to be the more reliable.

Similar conclusions could be obtained when considering the entire set of studied phenotypes. The strains, grouped either by isolation source or genotype, did not show significant differences in the overall PCoA coordinates, confirming the lack of specific characteristics of wasp strains.

The reduced number of correlations among traits observed in the wasp strains compared to grape and wine strains could be ascribed to a high variance of phenotypes borne by strains isolated from wasp intestines. Indeed, the existence of correlations among traits relies in two points: (a) similar or opposite trends of traits; (b) homogeneity of the trends among the tested strains. Wine and grape strains showed several correlations among phenotypes, which could be observed only if the vast majority of the grouped strains bear similar characteristics. In addition, the presence of some correlations among traits in the wine group can be related to the selection of the strains for wine-making purposes: several volatile compounds are correlated with a strain's ability to grow at low temperatures, both traits useful for wine fermentations.

Taken together, our results indicate that wasp strains, in addition to representing a wide range of *S. cerevisiae* genotypes, also represent a large part of the phenotypes of the sympatric set of yeast strains. Our findings clearly indicate that strains isolated from wasp intestines represent a great resource of yeast phenotype biodiversity that should be preserved, to safeguard not only the natural variance of this microorganism but also the interests of wine-makers, who could take strong advantage from exploitation of their phenotypic variability.

Acknowledgements

The authors thank Centro di Servizi di Spettromeria di Massa (University of Florence, Florence, Italy), in particular G. Pieraccini, L. Turbanti and G. Moneti, for their support in metabolic data acquisition and analysis. This project was supported by SYBARIS (Grant No. 242220). Supplementary funds were provided by the University of Florence.

References

Borcard D, Gillet F, Legendre P. 2011. Numerical Ecology with R. Springer: New York.

Callejon RM, Clavijo A, Ortigueira P, *et al.* 2010. Volatile and sensory profile of organic red wines produced by different selected autochthonous and commercial *Saccharomyces cerevisiae* strains. *Anal Chim Acta* **660**: 68–75.

Christiaens JF, Franco LM, Cools TL, *et al.* 2014. The fungal aroma gene *ATF1* promotes dispersal of yeast cells through insect vectors. *Cell Rep* **9**: 1–8.

Fay JC, Benavides JA. 2005. Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet* **1**: 66–71.

Francesca N, Canale DE, Settanni L, *et al.* 2012. Dissemination of wine-related yeasts by migratory birds. *Environ Microbiol Rep* **4**: 105–112.

Goddard MR, Anfang N, Tang R, *et al.* 2010. A distinct population of *Saccharomyces cerevisiae* in New Zealand: evidence for local dispersal by insects and human-aided global dispersal in oak barrels. *Environ Microbiol* **12**: 63–73.

Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *J Statist Software* **22**: 1–19.

Hampsey M. 1997. A review of phenotypes in *Saccharomyces cerevisiae*. *Yeast* **13**: 1099–1133.

Hyma KE, Fay JC. 2013. Mixing of vineyard and oak-tree ecotypes of *Saccharomyces cerevisiae* in North American vineyards. *Mol Ecol* **22**: 2917–2930.

Hyma KE, Saerens SM, Verstrepen KJ, Fay JC. 2011. Divergence in wine characteristics produced by wild and domesticated strains of *Saccharomyces cerevisiae*. *FEMS Yeast Res* **11**: 540–551.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and

multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.
Knight S, Klaere S, Fedrizzi B, Goddard MR. 2015. Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir. *Sci Rep* **5**: 14233.
Knight S, Goddard MR. 2015. Quantifying separation and similarity in a *Saccharomyces cerevisiae* metapopulation. *ISME J* **9**: 361–370.
Le Cao KA, Gonzalez I, Dejean S. 2009. integrOmics: an R package to unravel relationships between two omics datasets. *Bioinformatics* **25**: 2855–2856.
Legras JL, Merdinoglu D, Cornuet JM, Karst F. 2007. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* **16**: 2091–2102.
Legras JL, Ruh O, Merdinoglu D, Karst F. 2005. Selection of hypervariable microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains. *Int J Food Microbiol* **102**: 73–83.
Liti G, Carter DM, Moses AM, *et al.* 2009. Population genomics of domestic and wild yeasts. *Nature* **458**: 337–341.
Liti G. 2015. The fascinating and secret wild life of the budding yeast *S. cerevisiae*. *Elife* **4**: e05835.
McGovern PE, Zhang J, Tang J, *et al.* 2004. Fermented beverages of pre- and proto-historic China. *Proc Natl Acad Sci U S A* **101**: 17593–17598.
Muller-Thurgau L. 1896. Uber den Ursprung der Weinhefe und Hieran sich knuepfende praktische Folgerungen. *Weinbau Weinhandel* **7**: 40–41.
Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
Radford A. 1991. Looking at Yeast Cells. In Methods in Yeast Genetics: A Laboratory Course Manual, Rose M, Winston F, Hieter P (eds). Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York.
Ramazzotti M, Berna L, Stefanini I, Cavalieri D. 2012. A computational pipeline to discover highly phylogenetically informative genes in sequenced genomes: application to *Saccharomyces cerevisiae* natural strains. *Nucleic Acids Res* **40**: 3834–3848.
Robinson J. 2006. Chianti. The Oxford Companion to Wine. Oxford University Press: Oxford.
Romano P, Monteleone E, Paraggio M, *et al.* 1998. A methodological approach to the selection of *Saccharomyces cerevisiae* wine strains. *Food Technol Biotechnol* **36**: 69–74.
Romano P, Suzzi G, Mortimer R, Polsinelli M. 1995. Production of high levels of acetoin in *Saccharomyces cerevisiae* wine yeasts is a recessive trait. *J Appl Bacteriol* **78**: 169–174.
Stefanini I, Dapporto L, Legras JL, *et al.* 2012. Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution. *Proc Natl Acad Sci U S A* **109**: 13398–13403.
Styger G, Prior B, Bauer FF. 2011. Wine flavor and aroma. *J Ind Microbiol Biotechnol* **38**: 1145–1159.
Vopalenska *et al*. 2005. The morphology of *Saccharomyces cerevisiae* colonies is affected by cell adhesion and the budding pattern. *Res Microbiol* **156**: 921–931.
Wang QM, Liu WQ, Liti G, *et al.* 2012. Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. *Mol Ecol* **21**: 5404–5417.
Warringer J, Zorgo E, Cubillos FA, *et al.* 2011. Trait variation in yeast is defined by population history. *PLoS Genet* **7**e1002111.
Wayne RK, Ostrander EA. 2007. Lessons learned from the dog genome. *Trends Genet* **23**: 557–567.
Wondra M, Berovic M. 2001. Analyses of aroma components of Chardonnay wine fermented by different yeast strains. *Food Technol Biotechnol* **39**: 141–148.
Yu J, Pressoir G, Briggs WH, *et al.* 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* **38**: 203–208.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site.

Figure S1 *Saccharomyces cerevisiae* population ancestries estimated by using the model-based program STRUCTURE (Pritchard et al. 2000) on the SNP sequences of three genes able to recapitulate the whole genome phylogeny (Ramazzotti et al. 2012).
Figure S2 Volatile compounds measured after red (right) and white must fermentation by single strains.
Figure S3 PLSDA on volatile compounds amount produced during fermentation of red and white must not corrected for population structure.
Figure S4 Comparison of cases (replicate ferments by single strains) distribution in PLSDA coordinates according to must type, strain isolation source and strains' genotype.
Figure S5 NMMDS analysis on the separated sets of phenotypes.
Figure S6 Mantel tests for univariate correlations among genetic and phenotypic dissimilarity matrices.
Figure S7 Survey of the fermentation mass loss data for the recording days.
Figure S8 Strains' growth abilities.
Figure S9 PCoA on growth ability traits.
Figure S10 Strains' resistance to stresses.
Figure S11 PCoA on resistance to stress traits.
Figure S12 PCoA on all phenotypic traits.
Table S1 Statistical analyses comparing the strains distribution according to the first two components of the PLSDA carried out on the metabolic profiles of ferments.
Table S2 Statistical analyses comparing the strains distribution according to the first two coordinates of the PCoA carried out on all phenotypic traits.