# Targeted $^{1}$-H-NMR wine analyses revealed specific metabolomic signatures of yeast populations belonging to the *Saccharomyces* genus.

Vion Charlotte[1,2], Le Mao Inès[2], Nadine Yeramian[3], Maïtena Muro[1,2], Margaux Bernard[1,2], Da Costa Gregory[2], Richard Tristan[2], Marullo Philippe[1,2]

[1]Biolaffort, Bordeaux, FRANCE

[2] UMR 1366 Œnologie, Université de Bordeaux, INRAE, Bordeaux INP, BSA, ISVV

[3]Microbiology Division, Department of Biotechnology and Food Science, Faculty of Science-University of Burgos, Spain

Corresponding author:

## Abstract

This study aimed to explore the non-volatile metabolomic variability of a large panel of strains (44) belonging to the *Saccharomyces cerevisiae* and *Saccharomyces uvarum* species in the context of the wine alcoholic fermentation. For the *S. cerevisiae* strains *flor*, fruit and wine strains isolated from different anthropic niches were compared. This phenotypic survey was achieved with a special focus on acidity management by using natural grape juices showing opposite level of acidity. A $^{1}$H NMR based metabolomics approach was developed for quantifying fifteen wine metabolites that showed important quantitative variability within the strains. Thanks to the robustness of the assay and the low amount of sample required, this tool is relevant for the analysis of the metabolomic profile of numerous wines. The *S. cerevisiae* and *S. uvarum* species displayed significant differences for malic, succinic, and pyruvic acids, as well as for glycerol and 2,3-butanediol production. As expected, *S. uvarum* showed weaker fermentation fitness but interesting acidifying properties. The three groups of *S. cerevisiae* strains showed different metabolic profiles mostly related to their production and consumption of organic acids. More specifically, *flor* yeast consumed more malic acid and produced more acetic acid than the other *S. cerevisiae* strains which was never reported before. These features might be linked to the ability of *flor* yeasts to shift their metabolism during wine oxidation.

## Key words

$^{1}$H-NMR, fermenting yeast, wine metabolome, malic acid, *flor* and wine yeast strain.

## 1. Introduction

Metabolic activities of microorganisms strongly impact the chemical composition of fermented products and modify their nutritional and organoleptic properties [1,2]. Thus, the quantification of chemical compounds in fermented goods is a critical step to understand the role of microbes and to control their development. In this context, wine alcoholic fermentation has been particularly well investigated, and many wine compounds were identified using GC-MS, LC-MS and NMR [3–6]. As widely reviewed [7–9], the yeast metabolism impacts the chemical composition of resulting wines with direct enological consequences. Although they broadly share the same metabolic pathways, yeast species and strains show a high variability in their metabolites production. Thus, yeast species involved in grape juice fermentation are characterized by specific metabolic signatures for primary metabolites [10,11], fermentative esters [12–14], and off-flavor compounds [15]. In mixed cultures, metabolic profiling is also useful for characterizing microbial interactions [16–18] that may impact wine complexity [19]. At the intra specific level, analytical chemistry methods are also decisive for driving yeast selection aiming to better control ethanol content [20], wine acidity [10,21], volatile thiols content [22] or fermentative esters production [13].

The natural genetic variability of the species *Saccharomyces cerevisiae* [23] constitutes an important source of metabolomic variability that has been deciphered for volatile [24,25] and non-volatile compounds [26,27] by applying quantitative genetic. These studies required the use of analytical methods such as HPLC [28], enzymatic assays [26], targeted GC-MS [24] allowing the quantification, for large sample sizes of wine metabolites belonging to different chemical family. Alternatively, untargeted metabolomics approaches may be applied for quantifying hundreds of metabolic features able to discriminate the biochemical signature of few *Saccharomyces cerevisiae* strains during the alcoholic fermentation [29][5]. However, the exact identification of such discriminating compounds remains a chemical challenge [30].

In this context, the exploration of extracellular metabolomic variability of yeast strains during the alcoholic fermentation for different classes of compounds constitute a challenging task. A performing and versatile analytical technique to quantify yeast metabolites is the proton Nuclear Magnetic Resonance ($^1$H NMR) spectroscopy [31,32]. The main $^1$H NMR spectroscopy assets are the simplicity of sample preparation and high reproducibility of the quantification. It is also possible to investigate a large range of metabolites belonging to different chemical families that are displayed on the same spectrum. This non-destructive technique also gives the opportunity to carry out several analyses on the same sample. NMR spectroscopy is quantitative since the signal intensity is directly proportional to the metabolite concentration and the number of nuclei in the molecule [33,34]. Despite its high potentiality to study fermented products, $^1$H NMR investigations are still scarce and face two major challenges [31]. First, the sensitivity of the technique is lower than mass spectrometry which explains that the latter is often preferred for non-targeted analyses. This issue can be addressed by applying a high number of

scans during the analysis, but this will extend the time required for the analysis. Secondly, peak overlaps from multiple metabolites pose major challenges. 2D NMR or supplementation with pure compounds of interest can be carried out to address this issue [35].

In the present study, we explored the metabolomic variability of different wine yeasts with a special focus on the management of wine acidity by malic acid. Indeed, the level of this organic acid tends to drastically decrease in grape juices due to climate change with important enological consequences [36]. As previously reported by several authors, strains and species of the *Saccharomyces* genus may modulate malic acid concentrations of wines [37–43]. Recently, pools of alleles belonging to *flor* or *wine* yeast populations were partially linked to the metabolic variability of malic acid [44] suggesting a possible effect of the yeast ecological origin. In this context, we reevaluated the natural variability of several *Saccharomyces* strains presenting different level of malic acid production or consumption during wine fermentation. In order to have a wide overview of possible metabolic trade-off related to malic acid metabolism, we developed new protocols aiming to quantify extracellular yeast metabolites using $^{1}$H-NMR based metabolomics, adapting methods already used for describing wine composition [45, 46].

## 2. Materials and Methods

### 2.1. Yeast strains used and culture methods

The forty-four strains of *S. cerevisiae* and *S. uvarum* isolated in different enological niches are listed in Table 1. *S. cerevisiae* strains were propagated on YPD 2 % (1 % peptone, 1 % yeast extract, 2 % glucose) at 30°C in both liquid and plate cultures (2 % agar). *S. uvarum* strains were propagated on YPD 6 % to avoid sporulation as described by [47]. Long term storage at -80°C was achieved by adding one volume of glycerol to YPD overnight cultures.

Table 1: Yeast strains used.

| Strain | Species | Ecological niche | Substrate/origin | Geographical area | Collection | Reference |
|---|---|---|---|---|---|---|
| FMGS1_889 | *S. cerevisiae* | Control | Extreme malic acid consuming strain | Miscelaneous (Breeding) | UMR Oenology (ISVV) | [21] |
| FMGS3_191 | *S. cerevisiae* | Control | Extreme malic acid producing strain | Miscelaneous (Breeding) | UMR Oenology (ISVV) | [70] |
| 14.280 (SV2) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 23.10 (S34V) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 34.220 (SV3) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 36.2J (S4V) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 8.1J (S6V) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| DBVPG4695 | *S. cerevisiae* | Flor | Vino Santo | Lungarotti winery, Italy | 1002 genomes | [23] |
| CBS4079 | *S. cerevisiae* | Flor | Velum | Spain | 1002 genomes | [23] |
| CBS4092 | *S. cerevisiae* | Flor | Velum | Spain | 1002 genomes | [23] |
| CBS4093a | *S. cerevisiae* | Flor | Velum | Spain | 1002 genomes | [23] |
| Y-1301 | *S. cerevisiae* | Flor | Wine | Unknown | NRRL collection | |
| YB-210 | *S. cerevisiae* | Fruit | Spoiled banana | Costa Rica | NRRL collection | |
| Y-747 | *S. cerevisiae* | Fruit | Cider | Illinois, USA | NRRL collection | |
| YB-1191 | *S. cerevisiae* | Fruit | Citrus juice | Louisiana, USA | NRRL collection | |
| Y-6678 | *S. cerevisiae* | Fruit | Olives | Spain | NRRL collection | |
| Y-2230 | *S. cerevisiae* | Fruit | Fruit juice | The Netherlands | NRRL collection | |
| YB-360 | *S. cerevisiae* | Fruit | Applesauce | Unknown | NRRL collection | |
| YB-2541 | *S. cerevisiae* | Fruit | Benzolated cider at 18C | Unknown | NRRL collection | |
| Y-641 | *S. cerevisiae* | Fruit | Cider | Unknown | NRRL collection | |
| Y-35 | *S. cerevisiae* | Fruit | Fruit (Ilex aquifolium) | Unknown | NRRL collection | |
| Y-6275 | *S. cerevisiae* | Fruit | Orange concentrate | Unknown | NRRL collection | |
| Y-129 | *S. cerevisiae* | Fruit | Orange juice | Unknown | NRRL collection | |
| YB-4081 | *S. cerevisiae* | Fruit | Ripe goyave | Unknown | NRRL collection | |
| YB-2573 | *S. cerevisiae* | Fruit | Sauerkraut | Unknown | NRRL collection | |
| YB-369 | *S. cerevisiae* | Fruit | Sauerkraut | Unknown | NRRL collection | |
| Y-964 | *S. cerevisiae* | Fruit | Sour figs | Unknown | NRRL collection | |
| Y-767 | *S. cerevisiae* | Fruit | Tomato product (B-117) | Unknown | NRRL collection | |
| RC4-15 | *S. cerevisiae* | Uvarum | Wine | Alsace, France | UMR Oenology (ISVV) | [49] |
| BR6-2 | *S. uvarum* | Uvarum | Cider | Britany/Normandy, France | UMR Oenology (ISVV) | [49] |
| CBS 377 | *S. uvarum* | Uvarum | Fruit juice | Germany | UMR Oenology (ISVV) | [49] |
| P3 | *S. uvarum* | Uvarum | Wine | Sancerre, France | UMR Oenology (ISVV) | [49] |
| CBS 425 | *S. uvarum* | Uvarum | Cider | Switzerland | UMR Oenology (ISVV) | [49] |
| GN | *S. cerevisiae* | Wine | meiotic spore clone from Zymaflore VL1 | Bordeaux, France | UMR Oenology (ISVV) | [43] |

Formatted Table

| | | | | | | |
|---|---|---|---|---|---|---|
| C1-4 | *S. cerevisiae* | Wine | Wine | Cordoba, Spain | UMR Oenology (ISVV) | [40] |
| C4-2 | *S. cerevisiae* | Wine | Wine | Cordoba, Spain | UMR Oenology (ISVV) | [40] |
| C9-10 | *S. cerevisiae* | Wine | Wine | Cordoba, Spain | UMR Oenology (ISVV) | [40] |
| SB | *S. cerevisiae* | Wine | meiotic spore clone from Actiflore BO213 | France | UMR Oenology (ISVV) | [43] |
| F15msp | *S. cerevisiae* | Wine | meiotic spore clone from Zymaflore F15 | France | UMR Oenology (ISVV) | [72] |
| M10-7 | *S. cerevisiae* | Wine | Wine | Madrid, Spain | UMR Oenology (ISVV) | [40] |
| M2-2 | *S. cerevisiae* | Wine | Wine | Madrid, Spain | UMR Oenology (ISVV) | [40] |
| M2-9 | *S. cerevisiae* | Wine | Wine | Madrid, Spain | UMR Oenology (ISVV) | [40] |
| M2msp | *S. cerevisiae* | Wine | meiotic spore clone from Enoferm M2 (Lallemand) | Unknown | UMR Oenology (ISVV) | [72] |
| BO213 | *S. cerevisiae* | Wine | Wine starter | France | Laffort | [26] |

Formatted Table

Formatted Table

Formatted Table

Formatted Table

Formatted Table

### 2.1.1. Grape Juices

Two grape juices, Sauvignon blanc 2019 (SB19) and Grenache 2021 (GR21) were collected in Bordeaux and Montpellier areas, respectively and were stored at -20°C. The GR21 is a red grape juice representing warm climate matrix with a very low malic acid content. The SB19 is a white grape juice that was supplemented with L-malic acid in order to artificially increase its acidity. The final composition of these grape juices in fermenting sugar, malic acid content and pH is listed in Table 2.

Table 2: Composition of the different grape juices used in the experiments.

| Composition | SB19 [a] | GR21 |
|---|---|---|
| L-malic acid (g/L) | 5.31 | 0.52 |
| pH | 3.20 | 3.52 |
| Fermentable sugars (g/L) | 202.4 | 240.0 |
| Assimilable nitrogen (mg N/L) | 124 | 106 |
| Total $SO_2$ (mg/L) | 53 | 32 |

Formatted Table

*[a] The SB19 has been supplemented to increase its malic acid content and decrease the pH. Its original composition was a malic acid concentration at 3.73 g/L and a pH of 3.38.*

### 2.1.2. Alcoholic fermentation monitoring

Small-volume alcoholic fermentations were implemented in screwed vials fermentations according to the general procedure described in [48]. Rapidly, 20 mL-screwed vials (Thermo Fisher Scientific, Bordeaux, France) were filled with 12 mL of grape juice and were tightly closed with screw cap-

magnetic (Agilent Technologies, hdsp cap 18 mm PTFE/il 100 pk, Les Ulis, France) perforated with hypodermic needles (G26-0.45 x 13 mm, Terumo, Shibuya, Tokyo, Japan) for allowing $CO_2$ release. Vessel was inoculated by $2.10^6$ viable cell.$mL^{-1}$ precultured in liquid media 50 % filtered must, 50 % sterile $H_2O$ for 24h. Cellular concentration and viability was estimated by flow cytometry using a CytoFlex (Beckman Coulter, Villepinte, France). Fermentation took place at 24°C in shaken vials by using an orbital shaker (SSL1, Stuart, Vernon Hills, IL, USA) at 175 rpm. Fermentation kinetics were estimated by monitoring manually (1-2 times per day) the weight loss caused by $CO_2$ release using a precision balance with automatic weight recording (LabX system, Mettler Toledo, Viroflay, France). The amount of $CO_2$ released according to time was modeled by local polynomial regression fit [48]. This model allows the estimation of the time necessary to reach the maximum $CO_2$ produced, the lag phase, the speed between 15 % and 50 % of the fermentation (V15_50), the speed between 50 % and 80 % (V50_80), the time to reach 80 % of $CO_2$ produced (T80) and the maximal theoretic $CO_2$ produced ($CO_2$max). Final pH was monitored using Five Easy Plus pH-meter (Mettler Toledo, Viroflay France) with a micro probe LE422 (Mettler Toledo, Viroflay France).

### 2.2. Microsatellite genotyping

The genomic DNA of the 39 *S. cerevisiae* strains was quickly extracted in 96-wells microplate format using a customized LiAc-SDS protocol [49]. Fourteen polymorphic microsatellite loci (*SCAAT3*, *C3*, *C5*, *SCYOR267C*, *C8*, *C11*, *SCAAT2*, *YKL172*, *C9*, *C4*, *SCAAT5*, *SCAAT1*, *C6*, *SCAAT5*, *SCAAT1*, *C6*, *YPL009*) were used for estimating the genetic relationships within those strains using PCR conditions previously described [50]. Two multiplex PCRs allowing genotyping of seven loci were carried out in a final volume of 12.5 µL containing 6.25 µL of Qiagen Multiplex PCR master mix and 1 µL of DNA template. 1.94 µL of each mix was added in the mixture using the concentrations indicated. Both reactions were run with the following program: initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 57°C for 2 min, 72°C for 1 min, and a final extension at 60°C for 30 min. The size of PCR products was analyzed by the MWG company (Ebersberg, Germany) using 0.2 µL of 600 LIZ (GeneScan) as a standard marker. Chromatograms were analyzed with the GeneMarker (V2.4.0, Demo) program.

### 2.3. Metabolites analysis by $^1$H NMR

#### *2.3.1. Wine samples preparation*

Samples were collected during the alcoholic fermentation. Sixty µL of centrifuged wine sample was diluted ten times in phosphate buffer prepared in $D_2O$ (pD 4.2, 0.1 M final). Sixty µL of solution of calcium formate (FCa) 1 mM and trimethylsilylpropanoic acid (TSP) 1.25 mM in $D_2O$ was added to the preparation. The FCa is used as an internal standard for calculating the concentration of metabolites and the TSP to set the spectrum at 0 ppm.

### 2.3.2. NMR spectra acquisition

Spectra were recorded on a 600 MHz Avance III NMR spectrometer (Bruker, Wissembourg, France) operating at 600.25 MHz, equipped with a TXI 5 mm probe with z gradient coils. The measurement was performed at 293 K using TopSpin 4.0.8 software (Bruker, Wissembourg, France). ~~The~~A 1D-NOESY pulse sequence (*noesygppr1d*) was used ~~was a *noesysygpps1d*. The acquisition~~with low power presaturation at the water frequency during relaxation delay and mixing time and spoil gradient. Relevant parameters were ~~set as follows: 64 free induction decays (FIDs) were collected into a time domain~~: spectral size: 64k; number of ~~64k data points, with a~~scans: 64; spectral width ~~of 16~~: 18 ppm~~,~~ ~~an~~; acquisition time ~~of~~: 3.4064 s ~~and a~~; relaxation delay ~~(RD) of~~: recovery delay: 5 s ~~per scan~~; mixing time: 100 ms. The 90° pulse calibration was carried out for each sample automatically, and the shimming was set manually in *gs* mode for each spectrum in order to obtain the finest possible line width (lower than 1 Hz). ~~Water suppression was achieved during the RD using a shape pulse with a band selective solvent suppression (20 Hz centered on water signal), with a power level for presentation of 50.37 dB and a shaped pulse for presaturation of 34.83 dB.~~ The FIDs were multiplied by an exponential function corresponding to a 0.3 Hz line-broadening factor prior to the Fourier transformation. Manual phase followed by automatic baseline corrections were applied to the resulting spectrum, which was aligned to zero using the TSP signal.

Wine metabolite identification was performed using databases, literature data [51], and addition of pure standards when faced with uncertainties. Additionally, some ~~two-dimensional~~2D experiments were used (TOCSY and HSQC) to insure the identification. The FCa signal at 8.28 ppm was used as internal standard for calculating the concentration of the identified compounds. Peak deconvolutions were performed by the global spectral deconvolution method (GSD) [52], using the simple mixture analysis (SMA) plugin of MestReNova 12.0 software (Mestrelab Research, Santiago de Compostela, Spain). Quantification was achieved according to the formula of Goldelmann *et al.* [62].

## 2.4. Statistical analyses

All the statistical and graphical analyses were carried out using R software [53] and plots were generated using the base or *ggplot2* packages.

### 2.4.1. Phylogenic analysis

The microsatellite dataset was manipulated using the *adegenet* package implemented in R and the genetic distance within the strains was estimated using the Bruvo's distance using the *poppr* package [73]. The phylogram was built by Neighbor Joining (*ape* package) [74]. In order to evaluate a possible genetic structuration between *fruit*, *wine* and *flor* yeasts, a Discriminant Analysis of the seventh Principal Components was applied using the *adegenet* package omitting the two control strains (FMGS_889 and AC1_191). The pairwise *Fst* between populations was estimated using the *genet.dist* function.

### 2.4.2. Multivariate analyses

Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red

Spearman correlation tests between traits were computed using the *corr* function (*psych* package) corrected for multiple tests using the Benjamini-Hochberg method (α = 0.001). Results were displayed with the *corrplot* function (*corrplot* package). The multivariate phenotypic variability of yeast strains was visualized by a Principal Component Analysis (PCA) using the *ade4* package. In order to highlight the most discriminating variables of PCA, Discriminant Analysis of Principal Component (DAPC) (*adegenet* package) was applied by selecting principal components allowing to capture up to 60% of the cumulated inertia.

#### *2.4.3. Analysis of variance*

Analyses of Variance (ANOVA) were carried out using the *car* package. The phenotypic values measured in the grape juices SB19 and GR21 were analyzed using the linear model (LM1) in order to estimate the effect of the following factors: media, population, strain in population as well as the first order interaction of all the factors according to the formula (1).

(1) $y_{ijk} = media_i + population_j + strain(population)_{jk} + inter^2_{ijk} + \varepsilon_{ijkl}$

where *y* is value of all the variables for a media *i* (*i*=1,2) in which *j* (*j*=1,2,3,4) groups of yeast strains fermented. Each population is composed of *k* strains. The factor strain is nested in the factor population and *k* varies between 1:5 and 1:16 according to the number of strains per population. The term $inter^2_{ijk}$ represents the first order interaction of each factor and $\varepsilon_{ijkl}$ the residual. The analysis of variance of model LM1 allows the estimation of the primary effect of the media, population, and strain in population on several quantitative variables, as well as their primary interaction effect. The normal distribution of residues as well as the homoscedasticity of variances were tested by Shapiro test and Levene test (*car package*), respectively. When necessary, non-parametric comparison of samples were carried out using the Wilcoxon-Mann-Whitney or Kruskal test with corrected p values (Benjamini-Hochberg method, α = 0.05).

## 3. Results

### 3.1. Genetic characterization of a panel of *Saccharomyces* strains

The forty-two strains used were classified according to their origin and were denominated thereafter "*flor*" (n=10), "wine" (n=11), "fruit" (n=16), and "uvarum" (n=5) (Table 1). *Flor* and wine yeast populations (*S. cerevisiae*) share the same ecological niches but have been clearly separated by molecular phylogenetic studies [54,55]. As mentioned in the introduction, several genes related to central metabolism and proton homeostasis have been linked to the ability to consume, or not, malic acid during the alcoholic fermentation [44]. For most of them, alleles of the *flor* origin were related to a stronger consumption of malic acid suggesting a possible metabolic adaptation of this population to malic acid. However, phenotypic differences of malic acid consumption between *flor* and wine yeasts were never reported to our knowledge. To enlarge the genetic diversity of this metabolic survey, several *S. cerevisiae* strains (16) related to the fermentation of fruits (agrums, apple) and acidic substrates

(sauerkraut or tomato juice) were included in the panel. Finally, five *S. uvarum* strains presenting different genetic origin [42] were integrated to this panel since this psychrophilic species of the *Saccharomyces* genus is a strong producer of malic acid [56]. In addition, two *S. cerevisiae* strains AC1_191 and FMGS_889 were included to this panel as "control" for a total of 44 strains. These last two strains have been recently selected for their extreme ability to consume or produce malic acid [21,57] and have a mixed inheritance between *flor* and wine origin. The Bruvo's genetic distance between the 37 *Saccharomyces cerevisiae* strains used was computed by using 14 microsatellites markers as previously described [26]. The genetic distances range is between 0.11 and 0.96, indicating that all the strains are unique as illustrated by the genetic tree shown in Figure 1A with some confusions between *flor* and wine populations. A discriminant analysis of the ten first Principal Components (60 % of the cumulative variance) allows the separation of the three populations (Figure 1B) with a probability of assignation of 0.90, 1.00 and 0.93 for *flor*, wine, and fruit populations respectively. The pairwise Fst between *flor* and wine strains was 0.169 indicating a clear separation between the two populations. The five strains of *S. uvarum* are also unique and represent the variability of holoarctic *S. uvarum* as reported in a former study [42].

## 3.2. Assessment of wine metabolites variability of fermenting yeasts by targeted $^{1}$H-NMR analysis.

### *3.2.1. ~~Targeted~~ $^{1}$H-NMR analysis of yeast*

A targeted $^{1}$H-NMR metabolomics approach was applied for quantifying metabolites of fermenting yeasts belonging to the *Saccharomyces* genus. The main wine metabolites produced and consumed by yeasts during the alcoholic fermentation were quantified from sixty microliters of wine stored at -80°C. Samples were simply thawed and diluted in an appropriate buffer before analysis as detailed in methods. The typical $^{1}$H-NMR spectrum after water suppression is presented in Figure 2. The signals at 0.00 ppm and 8.28 ppm correspond to TSP and FCa respectively; other signals correspond to wine constituents. The $^{1}$H-NMR spectra were dominated by ethanol, and glycerol, followed by organic acids. Even if different amino acids were observed in juice, due to their consumption by yeast [58], only tyrosine and arginine were quantified after the alcoholic fermentation. The Table 3 shows the chemical shifts and the coupling constants used for identification and quantification of 15 metabolites including six organic acids (acetic acid, citric acid, malic acid, pyruvic acid, succinic acid, and tartaric acid), three alcohols (ethanol, 2,3-butanediol, glycerol), two reducing sugars (fructose and glucose), one ester (ethyl acetate), one aldehyde (ethanal), and two amino acids (tyrosine and arginine). The concentration range and the average coefficient of variation (CV) of each metabolite measured in one red (GR21) and one white grape juice (SB19) is also indicated, as well as the average concentration in both grape juices.

**Table 3:** Typical chemical shifts and coupling constants used for compounds identification after fermentation. The signals chosen for quantification are in bold. Variation coefficients and average concentration are displayed for all compounds in both media.

| Compound | δ1 H (Multiplicity, J in Hz, Assignment) | Average CV SB19 | Relative Average concentration SB19 (g/L) | Average CV GR21 | Relative Average concentration GR21 (g/L) |
|---|---|---|---|---|---|
| acetic acid | **2.08** (s, CH3) | 21.06 | 0.2786 | 20.27 | 0.3166 |
| arginine | **1.68** (m, CH2) | 8.67 | 0.0031 | 43.09 | 0.0074 |
| 2,3-butanediol | **1.13** (d, 6.2, 2CH3), | 15.84 | 0.4027 | 17.77 | 0.5089 |
| citric acid | **2.~~67~~69** (d, 15.6, CH2), | 7.49 | 0.4265 | 49.74 | 0.0953 |
| ethanal | **2.23** (d, 3.0, CH3), 9.79 (q, 2.85, CH) | 60.44 | 0.0023 | 34.15 | 0.0141 |
| ethanol | **1.17** (t, 7.2, CH3), 3.65 (q, CH2) | 7.89 | 59.6569 | 4.82 | 104.2476 |
| ethyl acetate | 1.26 (t, 7.2, CH3), **2.03** (s, CH3), 4.12 (q, CH2) | 8.48 | 0.7078 | 5.97 | 1.1418 |
| fructose | 3.97 (dd, 10.1, 3.5, CH), **4.01** (m, CH), 4.09 (dd, 12.8, 1.0, CH) | 37.38 | 1.3929 | 49.99 | 1.8364 |
| glucose | **4.65** (d, 7.9, CH), **5.23** (d, 3.6, CH) | 25.47 | 0.2182 | 32.71 | 0.1882 |
| glycerol | **3.55** (dd, 11.8 and 6.5, 2CH2), 3.77 (m, CH) | 5.50 | 10.1624 | 9.01 | 9.0169 |
| malic acid | 2.59 (dd, 16.3 and 7.0, CH), **2.~~79~~81** (dd, 16.3 and 4.5, CH), 4.36 (dd, CH) | 5.44 | 6.9769 | 12.28 | 1.2349 |
| pyruvic acid | **2.38** (s, CH3) | 18.14 | 0.0542 | 21.83 | 0.0341 |
| succinic acid | **2.~~54~~58** (s, 2CH2) | 10.81 | 1.8568 | 15.24 | 1.4810 |
| tartaric acid | **4.39** (s, 2CH) | 40.52 | 0.5950 | 15.31 | 1.1052 |
| tyrosine | **6.86** (m, 2CH), 7.17 (m, 2CH) | 16.34 | 0.0470 | 15.05 | 0.0517 |

Formatted Table

In both grape juices, most compounds had a CV lower than 30 % and in the SB19, six compounds had a CV under 10 % which shows the good repeatability and reliability of the analysis. Some compounds (residual sugars, ethanal, arginine and tartaric acid) were not accurately quantified due to their low concentrations and their instability to sample concentration (see discussion).

The relative concentrations presented in the table 3 allows the statistical comparisons between samples. For some key compounds (acetic acid, malic acid, succinic acid, and glycerol~~)~~), we quantified their absolute concentrations by applying correcting factors calculated ~~as described by~~using the standard addition method and applied using the formula of Goldelman et al. [62] (table 4). This absolute quantification was compared to enzymatic assay methods commonly used in enology [43]. As expected, strong correlations between enzymatic and $^{1}$H-NMR quantifications were found for all the metabolites (Spearman correlation analysis). The average CV of both methods were similar in the two grape juices, demonstrating that $^{1}$H-NMR assay was also very reproducible for such metabolites (Figure S1 panels A

and B). For acetic acid, malic acid, and succinic acid, the high correlations coefficient (rho > 0.82) indicated a good level of agreement between the two methods. Glycerol concentrations were also correlated despite a weaker Spearman's coefficient (rho = 0.52, p-value = $2.79x10^{-21}$). Linear regressions suggested that quantification by $^{1}$H-NMR of acetic acid and glycerol were slightly overestimated respect to enzymatic assays (slopes of 0.82 and 0.74, respectively). In contrast succinic acid concentration was slightly overestimated (slope of 1.22). However, both methods resulted in very similar quantifications and were in agreement with enological values expected.

**Table 4:** Correction factors and absolute average concentration measured for acetic, malic, succinic acids and glycerol in red and white wines.

| Compound | Correction factors | | Absolute average concentration SB19 (g/L) | Absolute average concentration GR21 (g/L) |
|---|---|---|---|---|
| | SB19 | GR21 | | |
| acetic acid | 0.94 | 1.17 | 0.26 | 0.37 |
| glycerol | 0.63 | 0.64 | 6.40 | 5.77 |
| succinic acid | 0.44 | 0.44 | 0.82 | 0.65 |
| malic acid | 0.57 | 0.68 | 3.97 | 0.84 |

Formatted Table

Formatted Table

### *3.2.2. Multivariate analysis of wine metabolites*

Fermentations were monitored daily by following weight loss, and six kinetic parameters (lag phase, V15_50, V50_80, T35, T80 and $CO_2$max) were extracted as previously reported [48]. The fifteen end-point metabolites were quantified for a panel of 44 strains of different origins (Table 1) in two grape juices (SB19 an GR21) fermented in triplicate. A Principal Component Analysis (PCA) was applied to average values in order to capture the overall variability of the 15 wine metabolites quantified by $^{1}$H-NMR (Figure 3). The final concentrations of wine metabolites are mostly structured by the grape juice nature and are clearly separated by the first component (40.3 % of inertia). Indeed, the SB19 was enriched in malic acid and citric acid compared to the GR21. In contrast, GR21 displayed a higher sugar content resulting in a higher production of ethanol, ethyl-acetate and ethanal. Beside this grape juice effect, *S. uvarum* and *S. cerevisiae* species are partially separated by the second component which captured 16.9 % of the total inertia. This axis is mostly correlated with glycerol and succinic acid content which are overproduced by *S. uvarum* strains. In addition, the presence of reducing sugars at the end of the alcoholic fermentation was detected for some S. *uvarum* and *fruit* strains which is also strongly related to the axis 2. In contrast, the quantification of the 15 metabolites did not allow a clear separation of the three *S. cerevisiae* populations in both matrices.

In order to identify metabolic signatures able to better discriminate yeast populations, the two grape juice datasets were analyzed separately using a Discriminant Analysis of Principal Component (DAPC). In the SB19, malic acid is the most discriminating compound followed by glycerol and succinic acid. The cumulated variability explained by these three metabolites on main linear discriminant axis is 98.3

%. As shown in Figure 4A, wines fermented by the *S. uvarum* species are significantly more concentrated in malic acid, glycerol, and succinic acid (Kruskal test, alpha < 0.05) than *S. cerevisiae*. These two species were also well discriminated in GR21 grape juice by three compounds: malic acid, succinic acid, and acetic acid which represent 64 %, 31 %, and 2 % of the discriminating inertia, respectively (Figure 4B). Beside this major species effect, significant differences of minor magnitudes were also observed within *S. cerevisiae* populations with a special emphasis for wine and flor populations that were always significantly different. In GR21 grape juice, the four populations were better separated than in the SB19 grape juice. Indeed, in GR21 each population was statistically different for malic acid content (Kruskal-Wallis' test, p-value < 0.05) while in SB19, only *S. uvarum* produced significantly more malic acid than the three other groups. This result could be linked to the initial characteristics of GR21 that contained a very low amount of malic acid.

### *3.2.3. Contributions of genetic and environmental effect on the phenotypic variability of Saccharomyces strains*

To deeply investigate factors influencing kinetic and metabolic traits, a nested analysis of variance was applied, aiming to estimate the impact of environment (grape juice) and genetic factors and their possible interactions. The genetic contribution effect was decomposed in population effect and strain within population effect as detailed by the linear model 1 (see methods). The contribution of each factor on 21 quantitative traits is summarized in Figure 5. The variability of ethanol, citric acid, malic acid, ethyl acetate, tartaric acid, and ethanal was mainly due to the grape juice effect. Indeed, ethanol and ethanal are directly linked to the initial sugars concentration that differs between red and white grape juices (SB19 = 202.4 g/L, GR21 = 240 g/L). Ethyl acetate is a carboxylate ester which is a secondary metabolite produced during alcoholic fermentation, derived from ethanol [59]. Even though its production can vary according to the strain, the very different ethanol content between the two wines led to significantly different amount of ethyl acetate. Tartaric acid is almost two times more concentrated in GR21 than SB19 (1.10 g/L vs 0.59 g/L) and the average amount of citric acid is about 4 times more important in SB19 than in GR21 (0.426 g/L vs 0.095 g/L respectively). Finally, malic acid varied greatly with the environment as the two grape juices were selected for their extreme acidities.

To better estimate the genetic contribution on malic acid content, we used the variable *MAC* (Malic Acid Consumed) that represents the ratio of malic acid consumed expressed in percentage (Figure 6A). *MAC* variability was evenly influenced by G (33 % var), E (36 % var) and GxE interactions (22 % var). The *MAC* values of the two control strains AC1_191 [57] and FMGS_889 [21] highlighted their great impact on acidity management since they produced and consumed significantly more than any other *S. cerevisiae* strains of this study. Quantitatively, the differences of absolute malic acid concentrations between wines made by these extreme strains were 1.71 and 2.95 g/L for GR21 and SB19, respectively. Intuitively, the *MAC* is strongly correlated to the final pH (Figure 6B) since wine pH depends on the

concentration of grape juice organic acids and especially the amount of malic acid. The Figure 6C shows the correlation of *MAC* with other metabolites in both grape juices. Overall, malic producer strains such as *S. uvarum* and AC1_191 produced more succinic acid than other.

For other metabolites (acetic acid, 2,3-butanediol, fructose, glycerol, pyruvic acid, succinic acid, and tyrosine) as well as kinetics parameters (V15_50, V50_80, T35 and T80), the genetic effect is greater than the grape juice's one. The decomposition of genetic effect in population and strain contribution indicated that within the same population, the metabolic variability of strains is generally stronger than the metabolic variability observed between groups (Figure 5). This is well illustrated by acetic acid variability that is more explained by the strain within populations (65 %) than by population effect itself (11 %). This can be explained for instance by some *flor* (CBS4079) and fruit (Y-2230, Y-6678) strains that overproduced acetic acid compared to other strains from their own populations (Figure S2). Finally, arginine was the unique metabolite that was neither impacted by grape juice or yeast strain, likely due to its low concentration level at the end of the fermentation.

## 4. Discussion

### 4.1. $^{1}$H NMR as an effective tool for analyzing wine microorganism metabolites.

Alcoholic beverages are complex matrices composed of several hundred volatile and non-volatile molecules that participate to the overall quality of the product. Therefore, analytical chemistry efforts aiming to characterize such complex matrices must be done for understanding and quantifying the role of microorganisms that participate to their elaboration. From a methodological viewpoint, trade-offs exist between the number of biological samples to analyze, the number of compounds assayed, the volume of sample required for their quantification, and the cost of the analysis. Generally, most of the microbiology studies lay emphasis on the number of biological sample and conditions analyzed in order to better understand the complex interactions existing between microbiological diversity and environmental conditions. This wide phenotypic characterization requires a reduction of fermentation volumes that allows a more efficient parallelization of culture conditions [43]. However, this strategy generally reduced the number of metabolites investigated by reducing the available sample volume. Thus, it is necessary to use high throughput methods such as enzymatic assay [43] or HPLC analyses [28]. Enzymatic assays are cost effective and can be easily robotized, however they require a specific assay per compounds with a quite limited number of metabolite available. HPLC analyses are more expensive and require the use of specific methods for quantifying organic acid and sugars [60], or nitrogen compounds [61].

NMR metabolomics is particularly amenable to detect compounds that are less tractable by liquid chromatography such as sugars, organic acids, alcohols polyols, and other highly polar compounds [33]. Those classes of compounds are well represented in wine reinforcing the interest of this technique. In this study, we used an analytical method able to quantify from a small volume (less than 100 µL) using

quantitative $^{1}$H-NMR spectroscopy. The advantages of NMR are the simplicity of sample preparation, measurement rapidity and the possibility to detect compounds belonging to different chemical families on one spectrum, in a single experiment. Previous studies were focused on the characterization of commercial wines and analysis of the cultivar, geographic origin, or vintage [45,46,64]. Our study focused on the characterization of wines fermented in the laboratory with different yeasts strains. The $^{1}$H-NMR method allowed the identification and quantification of 15 wine soluble metabolites of 264 biological samples requiring only 60 µL of wine for each experiment. This original approach allows studying the origin of many strains and their impact on various wine metabolites in different enological matrices.

The NMR analytical method was cross assessed by comparing the quantifications obtained by enzymatic assays for three organic acids (malate, succinate, and acetate) and for glycerol. Data obtained were accurately correlated for these metabolites and both methods had low coefficients of variation. By applying correction factors experimentally determined. For all the four compounds assayed the correlation between methods was very satisfactory and the slope of the linear models close to 1 +/- 20 % (Figure S1). For other compounds the absolute quantification was not determined but the relative concentrations showed in the table 3 did not impair the relative comparison of the strains which is the scope of this study. The quantification of glycerol seems more problematic even if the coefficient of variation of repetition is quite low (5.5% in SB19 and 9.01% in GR21). As shown Figure S1, the $^{1}$H-NMR method over-estimated glycerol content ~~possibly~~in comparison to enzymatic assays. The relative discrepancy between the two methods could be due to ~~the deconvolution of~~ several factors relating to both methods. Concerning NMR, the ~~double doublet at~~signal of glycerol ($\delta_H$ = 3.55 ppm ~~of $^{1}$H-NMR data. Indeed, this~~ ) is in a very dense region of the spectrum. This region is ~~very dense and~~ dominated by ethanol signal at 3.65 ppm. An overlap might have led to uncertainty of measurement.

In the present study, ethanol was quantified on all spectra as it is a compound of major interest in phenotypic analysis. However, the high intensity of the ethanol signal masks the signal of other compounds such as acids present in lower quantity (γ-aminobutyric acid, galacturonic acid, glucuronic acid, sorbic acid…). Therefore, the sensitivity of this method could be optimized in order to identify a higher number of compounds by suppressing the ethanol signal which is dominant after water using a selective pulse. Another solution would consist of introducing a lyophilization step to suppress the water and ethanol signals. This would improve the signal-to-noise ratio and better observe non-volatile compounds close to the signals of water and ethanol. These extra steps would increase the time of preparation and analysis but would allow the identification and accurate quantification of a higher number of other compounds.

Interestingly, the method applied is also useful for quantifying compounds that are not easily quantified by classical analyses such as ethyl acetate and ethanal. Ethyl acetate can be of interest as it is the major ester in several distilled spirits such as whiskey, rum and cachaça [63,64]. Ethanal, is found in wines in various concentrations and can be formed by yeasts or be the indication of a contamination

Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red
Formatted: Font color: Red

by acetic acid bacteria, lactic acid bacteria or be the result of auto-oxidation of ethanol and phenolic compounds [65]. The monitoring of its concentration in wines during fermentation and aging could be useful as its presence in small concentration gives a pleasant fruity aroma but at high concentrations it gives a pungent irritating odor. Finally, $^{1}$H-NMR analytical method is also applicable to grape juices and ongoing fermentations. The present study focused on the analysis of wines, but samples taken earlier during alcoholic fermentation could allow the assay of nitrogen compounds which are more abundant in the earliest stages of fermentation. On some grape juices' spectra, amino acids such as leucine, isoleucine, valine, threonine, arginine, proline, and choline were clearly observed (personal communication) in the $0.9 - 3.3$ ppm region of the spectrum. Due to their low concentration and consumption by yeast, these compounds were not quantified in fermented wines, except for arginine which is initially present in high concentration in grape juices and is lately consumed by yeasts [58].

### 4.2. Wine metabolome is partially impacted by the population origin of the fermenting strains.

The metabolomic characterization of two grape juices fermented by a large panel of yeast strains was achieved. Those strains provided by microorganism collections are associated with acidic fermented matrices such as wine, cider, fruit juices, and sauerkraut. As confirmed by microsatellite analysis, the four groups of strains collected in different ecological or human associated niches are discriminated by population genetics tools. The three *S. cerevisiae* groups belong to distinct populations (*flor*, *wine*, and *fruit*) the last group being constituted by *S. uvarum* strains isolated in cider and wine environments. This last species has been reported to have acidifying properties and to produce large quantities of malic acid [42].

In the context of this study, the impact of the yeast strain on wine metabolome was investigated in divergent matrices (one red and wine white grape juice) showing opposite level of malic acid content (0.52 g/L vs 5.31 g/L). Initially, the GR21 had a low content in malic acid compared to SB19 which enhanced the ability of some strains to produce more malic acid. As malic acid has a buffer effect, an addition of malic acid to the medium reinforced this buffer effect. On the contrary, in SB19 most strains consumed malic acid, limiting the buffer effect. The correlation between MAC and pH is strong in both grape juices (Figure 6B) as malic acid is the organic acid influencing the most the pH. Tartaric acid is also a strong influencer of the acidity, but it is not modulated by yeasts during alcoholic fermentation. Indeed, the figure 5 shows that this organic acid is only impacted by the grape juice and there is no significant effect of the strains or populations.

The main insight of this study is that metabolome variations are partially structured by the origin of the strain. First, the phenotypic discrepancies between the two main species involved in the alcoholic fermentation (*S. uvarum* and *S. cerevisiae*) were confirmed. Previous studies reported that *S. uvarum* produced higher concentrations of malic acid [42,11], succinic acid [66,67] and glycerol [11,67] that *S. cerevisiae*. In addition, this species is characterized by weaker fermentation performances than *S.*

448 *cerevisiae* [11,67] and by the production of specific fermentative aromas [66,68]. Da Silva *et al.* (2015)
449 reported that *S. uvarum* produced more acetic acid, but this finding was not supported in our study except
450 for one strain (RC4-15) which produced 0.70 g/L of acetic acid in GR21 while the 4 other *S. uvarum*
451 strains produced about 0.23 g/L in the same grape juice (figure S2). Using $^{1}$H-NMR-based
452 metabolomics, fifteen non-volatile metabolites were quantified. The results obtained confirmed the
453 differences reported for glycerol, malic acid, and succinic acid content (Figure 4). In addition, this
454 species reveals to produce more 2,3-butanediol (in average 1.5 times more than *S. cerevisiae* strains).
455 The amount of pyruvic acid was in average 2 times less important for *S. uvarum* than *S. cerevisiae*
456 strains, and residual fructose was the indication of some stuck fermentations by *S. uvarum* (~10 g/L of
457 fructose remaining).
458 Second, *wine* and *flor* yeasts can be distinguished by handful of compounds, mostly belonging to
459 organic acid family (cf. figure 4 and 5). This metabolic signature could be the consequence of a
460 differential management of central carbon metabolism as previously proposed by [44]. This might be
461 explained by the fact that *flor* yeasts are adapted to shift their metabolism to an oxidative metabolism
462 when sugar and nitrogen are depleted [69]. In the context of a shift diauxic during the velum formation
463 the aptitude to consume malic acid as a secondary carbon source may constitute a selective advantage.
464 Moreover, a correlation was found between malic acid production and succinic acid production in GR21
465 (no production occurred in SB19). This observation was previously reported by [70]. It is consistent
466 with their metabolic relationships and raise the question of the metabolic pathways taken by produced
467 and consumed malic acid. This correlation could explain a significative difference of succinic acid
468 content between *wine* and *flor* yeasts as they already differ for malic acid consumption.
469 Finally, the two control strains displayed extreme differences for malic acid content, demonstrating
470 the efficiency of the selection they came from [21,70]. They frame the natural variability of all the other
471 strains regarding the MAC values (Figure 6A). Only some strains of *S. uvarum* can produce as much
472 malic acid as AC1_191. On the PCA (Figure 3), the strain AC1_191 is close to *S. uvarum* group in terms
473 of acidity management. Indeed, this strain was selected for its ability to produce important amounts of
474 malic acid and happens to also produce important amount of succinic acid. The extreme strain AC1_191
475 appears to be a good strain to cope with low amounts of malic acid in musts in a context of climate
476 change. It is able to produce malic and, secondarily, succinic acid which decrease the pH of the final
477 wine [65]. Finally, it was able to complete the fermentation unlike some *S. uvarum* and produced low
478 amounts of acetic acid (Figure S3) which makes it suitable for winemaking. this strain could be used for
479 vinification itineraries aiming to conserve the freshness of white wines. The FMGS_889 is also a strain
480 of enological interest for lowering the acidity of rich malic wines and shortening the malo-lactic
481 fermentation of red wines [21]. The organoleptic consequences of the use of *S. cerevisiae* strains having
482 an opposite organic acid metabolism has been demonstrated in a previous study [57] and was not tested
483 in the present work due to the small volume of wine fermented.
484

## Conclusion

A reliable and easy to apply $^{1}$H-NMR analytical method was developed to quantify the major end point metabolites of the alcoholic fermentation of grape juices. This method required a small sample size (60 µL) a basic sample dilution and a short sample analysis (5 minutes). This efficient protocol was used for evaluating the metabolomic variability of 44 yeast strains becoming to the two major species of the *Saccharomyces* genus involved in wine fermentation. Metabolomic differences between strains belonging to different genetic groups were observed suggesting that the central metabolism of distinct populations is differently regulated. Interestingly, organic acids and glycerol metabolisms constitute strongly discriminating factors. The phenotypic diversity observed impacted the final pH value of wine and was strongly correlated to the ability of strains to consume or produce malic acid.


## Fundings

- « This study received financial support from the French government in the framework of the IdEX Bordeaux University "Investments for the Future" program / GPR Bordeaux Plant Sciences ». PM received a financial support from Aquitaine lander: program MEGAVIP, 2020-2023

## Acknowledgements

This work was supported by the Bordeaux Metabolome Facility and MetaboHUB (ANR-11-INBS-0010 project).

Authors warmly thanks Prof Hervé Alexandre, Prof Joseph Schacherer and Dr Jean-Luc Legras (CIRM) for providing some of the *flor* yeast strains used in this work.


## Author contribution

Conceptualization CV and PM; Formal analysis CV and PM; Funding acquisition PM; Investigation CV, MB, MM, GDC, ILM; Resources YN, PM; Software CV, PM; Supervision PM, TR; Writing – original draft CV, PM; Writing – review & editing CV, PM, GDC, TR

## References


1. Tamang JP, Watanabe K, Holzapfel WH. Review: Diversity of microorganisms in global fermented foods and beverages. Front Microbiol. 2016;7. doi:10.3389/fmicb.2016.00377
2. Bourdichon F, Casaregola S, Farrokh C, Frisvad JC, Gerds ML, Hammes WP, *et al.* Food fermentations: Microorganisms with technological beneficial use. Int J Food Microbiol. 2012;154: 87–97. doi:10.1016/j.ijfoodmicro.2011.12.030
3. Ebeler SE, Thorngate JH. Wine chemistry and flavor: Looking into the crystal glass. J Agric Food Chem. 2009;57: 8098–8108. doi:10.1021/jf9000555
4. Alañón ME, Pérez-Coello MS, Marina ML. Wine science in the metabolomics era. TrAC Trends Anal

Chem. 2015;74: 1–20. doi:https://doi.org/10.1016/j.trac.2015.05.006

5. Pinu FR. Grape and Wine Metabolomics to Develop New Insights Using Untargeted and Targeted Approaches. Fermentation. 2018. doi:10.3390/fermentation4040092

6. Lorrain B, Ky I, Pechamat L, Teissedre P-L. Evolution of Analysis of Polyhenols from Grapes, Wines, and Extracts. Molecules. 2013. pp. 1076–1100. doi:10.3390/molecules18011076

7. Fleet GH. Wine yeasts for the future. FEMS Yeast Res. 2008;8: 979–995. doi:10.1111/j.1567-1364.2008.00427.x

8. Ribéreau-Gayon, P., Dubourdien, D., Donéche B& AL. Handbook of Enology Volume 1 The Microbiology of Wine and Vinifications 2 nd Edition Handbook of Enology Volume 1 The Microbiology of Wine and Vinifications 2 nd Edition. 2006.

9. Swiegers JH, Bartowsky EJ, Henschke PA, Pretorius IS. Yeast and bacterial modulation of wine aroma and flavour. Aust J Grape Wine Res. 2005;11: 139–173. doi:10.1111/j.1755-0238.2005.tb00285.x

10. Hranilovic A, Gambetta JM, Schmidtke L, Boss PK, Grbin PR, Masneuf-Pomarede I, *et al*. Oenological traits of *Lachancea thermotolerans* show signs of domestication and allopatric differentiation. Sci Rep. 2018;8: 1–13. doi:10.1038/s41598-018-33105-7

11. Masneuf-Pomarède I, Bely M, Marullo P, Lonvaud-Funel A, Dubourdieu D. Reassessment of phenotypic traits for *Saccharomyces bayanus var. uvarum* wine yeast strains. Int J Food Microbiol. 2010;139: 79–86. doi:10.1016/j.ijfoodmicro.2010.01.038

12. Da Silva T, Albertin W, Dillmann C, Bely M, La Guerche S, Giraud C, *et al.* Hybridization within *Saccharomyces* genus results in homoeostasis and phenotypic novelty in winemaking conditions. PLoS One. 2015;10: 1–24. doi:10.1371/journal.pone.0123834

13. Viana F, Gil J V, Genovés S, Vallés S, Manzanares P. Rational selection of non-*Saccharomyces* wine yeasts for mixed starters based on ester formation and enological traits. Food Microbiol. 2008;25: 778–785. doi:https://doi.org/10.1016/j.fm.2008.04.015

14. Soles RM, Ough CS, Kunkee RE. Ester Concentration Differences in Wine Fermented by Various Species and Strains of Yeasts. Am J Enol Vitic. 1982;33: 94–98.

15. Chatonnet P, Dubourdie D, Boidron J, Pons M. The origin of ethylphenols in wines. J Sci Food Agric. 1992;60: 165–178. doi:https://doi.org/10.1002/jsfa.2740600205

16. Petitgonnet C, Klein GL, Roullier-Gall C, Schmitt-Kopplin P, Quintanilla-Casas B, Vichi S*, et al.* Influence of cell-cell contact between *L. thermotolerans* and *S. cerevisiae* on yeast interactions and the exo-metabolome. Food Microbiol. 2019;83: 122–133. doi:10.1016/j.fm.2019.05.005

17. Howell KS, Cozzolino D, Bartowsky EJ, Fleet GH, Henschke P a. Metabolic profiling as a tool for revealing *Saccharomyces* interactions during wine fermentation. FEMS Yeast Res. 2006;6: 91–101. doi:10.1111/j.1567-1364.2005.00010.x

18. Renault P, Coulon J, Moine V, Thibon C, Bely M. Enhanced 3-Sulfanylhexan-1-ol Production in Sequential Mixed Fermentation with *Torulaspora delbrueckii/Saccharomyces cerevisiae* Reveals a Situation of Synergistic Interaction between Two Industrial Strains. Frontiers in Microbiology. 2016. p. 293.

19. Tempère S, Marchal A, Barbe J-C, Bely M, Masneuf-Pomarede I, Marullo P, *et al*. The complexity of wine: clarifying the role of microorganisms. Appl Microbiol Biotechnol. 2018;102: 3995–4007. doi:10.1007/s00253-018-8914-8

Formatted: English (United States)

20. Tilloy V, Ortiz-Julien A, Dequin S. Reduction of ethanol yield and improvement of glycerol formation by adaptive evolution of the wine yeast *Saccharomyces cerevisiae* under hyperosmotic conditions. Appl Environ Microbiol. 2014;80: 2623–32. doi:10.1128/AEM.03710-13

21. Vion C, Peltier E, Bernard M, Muro M, Marullo P. Marker assisted selection of malic-consuming *Saccharomyces cerevisiae* strains for winemaking. Efficiency and limits of a qtl's driven breeding program. J Fungi. 2021;7. doi:10.3390/jof7040304

22. Dufour M, Zimmer A, Thibon C, Marullo P. Enhancement of volatile thiol release of *Saccharomyces cerevisiae* strains using molecular breeding. Appl Microbiol Biotechnol. 2013;97: 5893–5905. doi:10.1007/s00253-013-4739-7

23. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, *et al.* Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates Species-wide genetic and phenotypic diversity. Nature. 2018;556: 339–347.

24. Steyer D, Ambroset C, Brion C, Claudel P, Delobel P, Sanchez I, *et al*. QTL mapping of the production of wine aroma compounds by yeast. BMC Genomics. 2012;13: 573. doi:10.1186/1471-2164-13-573

25. Eder M, Sanchez I, Brice C, Camarasa C, Legras JL, Dequin S. QTL mapping of volatile compound production in *Saccharomyces cerevisiae* during alcoholic fermentation. BMC Genomics. 2018;19: 1–19. doi:10.1186/s12864-018-4562-8

26. Peltier E, Sharma V, Martí Raga M, Roncoroni M, Bernard M, Jiranek V, *et al.* Dissection of the molecular bases of genotype x environment interactions: A study of phenotypic plasticity of *Saccharomyces cerevisiae* in grape juices. BMC Genomics. 2018;19: 1–20. doi:10.1186/s12864-018-5145-4

27. Cubillos FA, Brice C, Molinet J, Tisné S, Abarca V, Tapia SM, *et al*. Identification of Nitrogen Consumption Genetic Variants in Yeast Through QTL Mapping and Bulk Segregant RNA-Seq Analyses. G3 Genes, Genomes, Genet. 2017;7: 1693–1705. doi:10.1534/g3.117.042127

28. Bony M, Bidart F, Camarasa C, Ansanay V, Dulau L, Barre P, *et al*. Metabolic analysis of *S. cerevisiae* strains engineered for malolactic fermentation. FEBS Lett. 1997;410: 452–456. doi:10.1016/S0014-5793(97)00637-6

29. Bordet F, Roullier-Gall C, Ballester J, Vichi S, Quintanilla-Casas B, Gougeon RD, *et al*. Different wines from different yeasts? "*Saccharomyces cerevisiae* intraspecies differentiation by metabolomic signature and sensory patterns in wine." Microorganisms. 2021;9: 1–17. doi:10.3390/microorganisms9112327

30. Reed LK, Baer CF, Edison AS. Considerations when choosing a genetic model organism for metabolomics studies. Curr Opin Chem Biol. 2017;36: 7–14. doi:10.1016/j.cbpa.2016.12.005

31. Cusano E, Cagliani LR, Consonni R, Simonato B, Zapparoli G. NMR-based metabolic profiling of different yeast fermented apple juices. Lwt. 2020;118. doi:10.1016/j.lwt.2019.108771

32. Avenoza A, Busto JH, Canal N, Peregrina JM. Time course of the evolution of malic and lactic acids in the alcoholic and malolactic fermentation of grape must by quantitative 1H NMR (qHNMR) spectroscopy. J Agric Food Chem. 2006;54: 4715–4720. doi:10.1021/jf060778p

33. Emwas AH, Roy R, McKay RT, Tenori L, Saccenti E, Nagana Gowda GA, *et al*. NMR Spectroscopy for Metabolomics Research. Metabolites. 2019;9. doi:10.3390/metabo9070123

34. Puig-Castellví F, Alfonso I, Piña B, Tauler R. 1H NMR metabolomic study of auxotrophic starvation in yeast using Multivariate Curve Resolution-Alternating Least Squares for Pathway Analysis. Sci Rep. 2016;6:

Formatted: English (United States)

Formatted: Spanish (Spain)

30982. doi:10.1038/srep30982

35. Vignoli A, Ghini V, Meoni G, Licari C, Takis PG, Tenori L, *et al.* High-Throughput Metabolomics by 1D NMR. Angew Chemie - Int Ed. 2019;58: 968–994. doi:10.1002/anie.201804736

36. Dequin S, Escudier J-L, Bely M, Noble J, Albertin W, Masneuf-Pomarède I, *et al.* How to adapt winemaking practices to modified grape composition under climate change conditions. OENO One. 2017;51: 205. doi:10.20870/oeno-one.2016.0.0.1584

37. Volschenk H, van Vuuren HJJ, Viljoen-Bloom M. Malic Acid in Wine: Origin, Function and Metabolism during Vinification. South African J Enol Vitic. 2006;27. doi:10.21548/27-2-1613

38. Van Leeuwen C, Destrac-Irvine A. Modified grape composition under climate change conditions requires adaptations in the vineyard. Oeno One. 2017;51: 147–154. doi:10.20870/oeno-one.2016.0.0.1647

39. Bigard A, Berhe DT, Maoddi E, Sire Y, Boursiquot JM, Ojeda H, *et al*. Vitis vinifera L. fruit diversity to breed varieties anticipating climate changes. Front Plant Sci. 2018;9: 0–16. doi:10.3389/fpls.2018.00455

40. Yéramian N, Chaya C, Suárez Lepe JA. L-(-)-malic acid production by *Saccharomyces* spp. during the alcoholic fermentation of wine. J Agric Food Chem. 2007;55: 912–919. doi:10.1021/jf061990w

41. Gao C, Fleet GH. Degradation of malic and tartaric acids by high density cell suspensions of wine yeasts. Food Microbiol. 1995.

42. Coloretti F, Zambonelli C, Castellari L, Tini V, Rainieri S. The effect of DL-malic acid on the metabolism of L-malic acid during wine alcoholic fermentation. Food Technol Biotechnol. 2002;40: 317–320.

43. Peltier E, Bernard M, Trujillo M, Barbe J-C, Prodhomme D, Gibon Y, *et al*. Wine yeast phenomics: A standardized fermentation method for assessing quantitative traits of *Saccharomyces cerevisiae* strains in enological conditions. PLoS One. 2018; 1–23. doi:10.1371/journal.pone.0190094

44. Peltier E, Vion C, Abou Saada O, Friedrich A, Schacherer J, Marullo P. Flor Yeasts Rewire the Central Carbon Metabolism During Wine Alcoholic Fermentation. Front Fungal Biol. 2021;2. doi:10.3389/ffunb.2021.733513

45. Le Mao I, Martin-Pernier J, Bautista C, Lacampagne S, Richard T, Da Costa G. 1H-Nmr Metabolomics as a Tool for Winemaking Monitoring. Molecules. 2021;26: 1–13. doi:10.3390/molecules26226771

46. Gougeon L, da Costa G, Guyon F, Richard T. 1H NMR metabolomics applied to Bordeaux red wines. Food Chem. 2019;301: 125257. doi:10.1016/j.foodchem.2019.125257

47. Albertin W, Da Silva T, Rigoulet M, Masneuf- Pomarede I, De Vienne D, Sicard D, *et al.* The Mitochondrial Genome Impacts Respiration but Not Fermentation in Interspecific *Saccharomyces* Hybrids. PLoS One. 2013.

48. Peltier E, Bernard M, Trujillo M, Prodhomme DD, Barbe J-C, Gibon Y, *et al*. Wine yeast phenomics: a standardized fermentation method for assessing quantitative traits of *Saccharomyces cerevisiae* strains in enological conditions. Schacherer J, editor. PLoS One. 2018;13: 191353. doi:10.1101/191353

49. Albertin W, Chernova M, Durrens P, Guichoux E, Sherman DJ, Masneuf-Pomarede I*, et al*. Many interspecific chromosomal introgressions are highly prevalent in Holarctic *Saccharomyces uvarum* strains found in human-related fermentations. Yeast. 2017;35: 141–156. doi:10.1002/yea.3248

50. Börlin M, Venet P, Claisse O, Salin F, Legras JL, Masneuf-Pomarede I. Cellar-associated *Saccharomyces cerevisiae* population structure revealed high-level diversity and perennial persistence at Sauternes wine estates. Appl Environ Microbiol. 2016;82: 2909–2918. doi:10.1128/AEM.03627-15

Formatted: English (United States)

642 51. Airoldi C, Tripodi F, Guzzi C, Nicastro R, Coccetti P. NMR analysis of budding yeast metabolomics: A
643 rapid method for sample preparation. Mol Biosyst. 2015;11: 379–383. doi:10.1039/c4mb00452c
644 52. Cobas C, Seoane F, Domínguez S, Sykora S. A new approach to improving automated analysis of proton
645 NMR spectra through Global Spectral Deconvolution (GSD). Spectrosc Eur. 2011;23: 26–30.
646 53. R Development Core Team. R: A language and environment for statistical computing. 2016;2: 1–12.
647 54. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A*, et al.* Genome evolution across
648 1,011 *Saccharomyces cerevisiae* isolates. Nature. 2018;556: 339–344. doi:10.1038/s41586-018-0030-5
649 55. Coi AL, Bigey F, Mallet S, Marsit S, Zara G, Gladieux P*, et al.* Genomic signatures of adaptation to wine
650 biological aging conditions in biofilm-forming flor yeasts. Mol Ecol. 2017. doi:10.1111/mec.14053
651 56. Chernova M, Albertin W, Durrens P, Guichoux E, Sherman DJ, Masneuf-Pomarede I, *et al*. Many
652 interspecific chromosomal introgressions are highly prevalent in Holarctic *Saccharomyces uvarum* strains found
653 in human-related fermentations. Yeast. 2018;35: 141–156. doi:10.1002/yea.3248
654 57. Vion C, Muro M, Bernard, Margaux, Richard B, Fautre V, Yeramian N, *et al.* New malic acid producer
655 strains of *Saccharomyces cerevisiae* for preserving wine acidity during alcoholic fermentation. Food Microbiol.
656 2022;112. doi:10.1016/j.fm.2022.104209
657 58. Crépin L, Nidelet T, Sanchez I, Dequin S, Camarasa C. Sequential use of nitrogen compounds by
658 *Saccharomyces cerevisiae* during wine fermentation: A model based on kinetic and regulation characteristics of
659 nitrogen permeases. Appl Environ Microbiol. 2012;78: 8102–8111. doi:10.1128/AEM.02294-12
660 59. Park YC, Shaffer CEH, Bennett GN. Microbial formation of esters. Appl Microbiol Biotechnol. 2009;85:
661 13–25. doi:10.1007/s00253-009-2170-x
662 60. Calull M, Marcé RM, Borrull F. Determination of carboxylic acids, sugars, glycerol and ethanol in wine
663 and grape must by ion-exchange high-performance liquid chromatography with refractive index detection. J
664 Chromatogr A. 1992;590: 215–222. doi:10.1016/0021-9673(92)85384-6
665 61. Lehtonen P. Determination of Amines and Amino Acids in Wine — A Review. Am J Enol Vitic. 1996;47:
666 127 LP – 133. doi:10.5344/ajev.1996.47.2.127
667 62. Godelmann R, Kost C, Patz CD, Ristow R, Wachter H. Quantitation of compounds in wine using 1H
668 NMR spectroscopy: Description of the method and collaborative study. J AOAC Int. 2016;99: 1295–1304.
669 doi:10.5740/jaoacint.15-0318
670 63. Nascimento ESP, Cardoso DR, Franco DW. Quantitative ester analysis in cachaça and distilled spirits by
671 Gas Chromatography-Mass Spectrometry (GC-MS). J Agric Food Chem. 2008;56: 5488–5493.
672 doi:10.1021/jf800551d
673 64. Alcarde AR, Souza LM, Bortoletto AM. Formation of volatile and maturation-related congeners during
674 the aging of sugarcane spirit in oak barrels. J Inst Brew. 2014;120: 529–536. doi:10.1002/jib.165
675 65. Liu SQ, Pilone GJ. An overview of formation and roles of acetaldehyde in winemaking with emphasis on
676 microbiological implications. Int J Food Sci Technol. 2000;35: 49–61. doi:10.1046/j.1365-2621.2000.00341.x
677 66. Pérez D, Jaehde I, Guillamón JM, Heras JM, Querol A. Screening of *Saccharomyces* strains for the
678 capacity to produce desirable fermentative compounds under the influence of different nitrogen sources in
679 synthetic wine fermentations. Food Microbiol. 2021;97. doi:10.1016/j.fm.2021.103763
680 67. Minebois R, Pérez-Torrado R, Querol A. A time course metabolism comparison among *Saccharomyces*
681 *cerevisiae, S. uvarum and S. kudriavzevii* species in wine fermentation. Food Microbiol. 2020;90: 103484.

doi:10.1016/j.fm.2020.103484

68. Pérez D, Denat M, Heras JM, Guillamón JM, Ferreira V, Querol A. Effect of non-wine *Saccharomyces* yeasts and bottle aging on the release and generation of aromas in semi-synthetic Tempranillo wines. Int J Food Microbiol. 2022;365. doi:10.1016/j.ijfoodmicro.2022.109554

69. Legras JL, Moreno-Garcia J, Zara S, Zara G, Garcia-Martinez T, Mauricio JC, *et al.* Flor yeast: New perspectives beyond wine aging. Front Microbiol. 2016;7: 1–11. doi:10.3389/fmicb.2016.00503

70. Vion C, Muro M, Bernard, Margaux, Richard B, Fautre V, Yeramian N, *et al.* New malic acid producer strains of *Saccharomyces cerevisiae* for preserving wine acidity during alcoholic fermentation. Food Microbiol. 2022.

71. David-Vaizant V, Alexandre H. Flor yeast diversity and dynamics in biologically aged wines. Front Microbiol. 2018;9: 1–16. doi:10.3389/fmicb.2018.02235

72. Huang C, Roncoroni M, Gardner RC. MET2 affects production of hydrogen sulfide during wine fermentation. Appl Microbiol Biotechnol. 2014;98: 7125–7135. doi:10.1007/s00253-014-5789-1

73. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: a R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014;2: e281. doi:10.7717/peerj.281

74. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 2004;20: 289–290. doi:10.1093/bioinformatics/btg412

Formatted: Spanish (Spain)

Formatted: English (United States)

## Figures

All figures should be printed in color

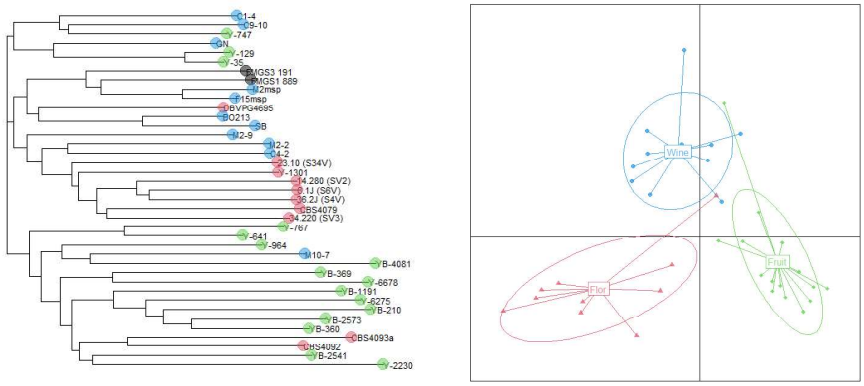
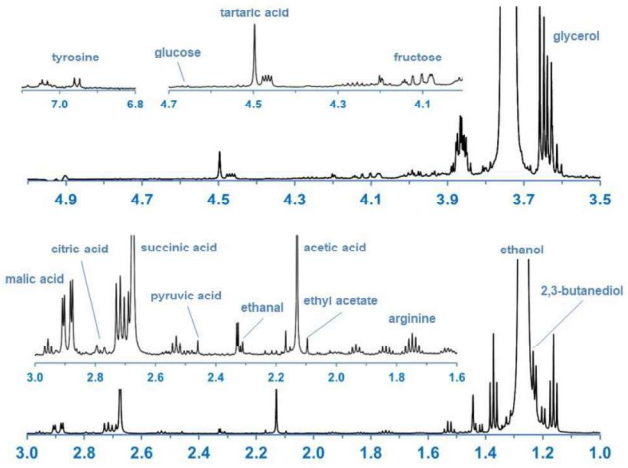
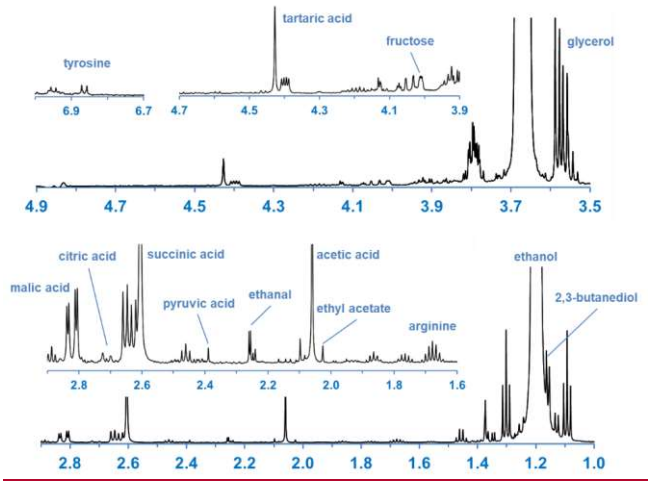


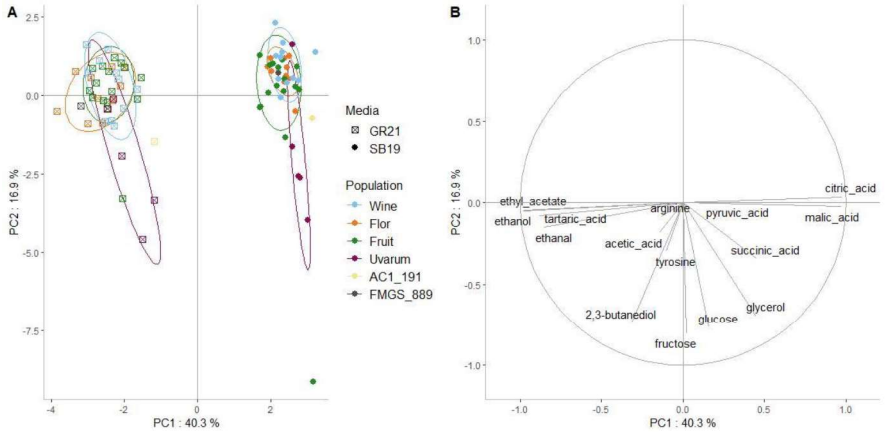

**Figure 1:** (A) Genetic relationships between the 37 S. cerevisiae strains of the study (Bruvo's distance) computed from the genotyping of 14 variable microsatellites, the colors red, green, and blue indicated the Flor, Fruit, and Wine origin, respectively. the two control strains are indicated by a black dot (b) Discriminant Analysis of Principal Components computed with microsatellite data. The three main populations are figured out by the same colors.






710

711 **Figure 2:** Annotated typical [1]H-NMR spectrum of wine metabolites after water suppression (~~NOESYGPPR1D~~noesygppr1d).
712 Identified constituents are listed in Table 3.




713

714 **Figure 3:** Principal Component Analysis (PCA) (A) The two first axis of the PCA performed for the 44 strains in the two grape
715 juices. Axes 1 and 2 explain 40.3 and 16.9 % of total variation respectively. Each point represents the average of three biological
716 replicates of a single strain. Each strain is colored according to its population. (B) Correlation circle indicating the corr elation
717 of the variables for axes 1 and 2.

718

719

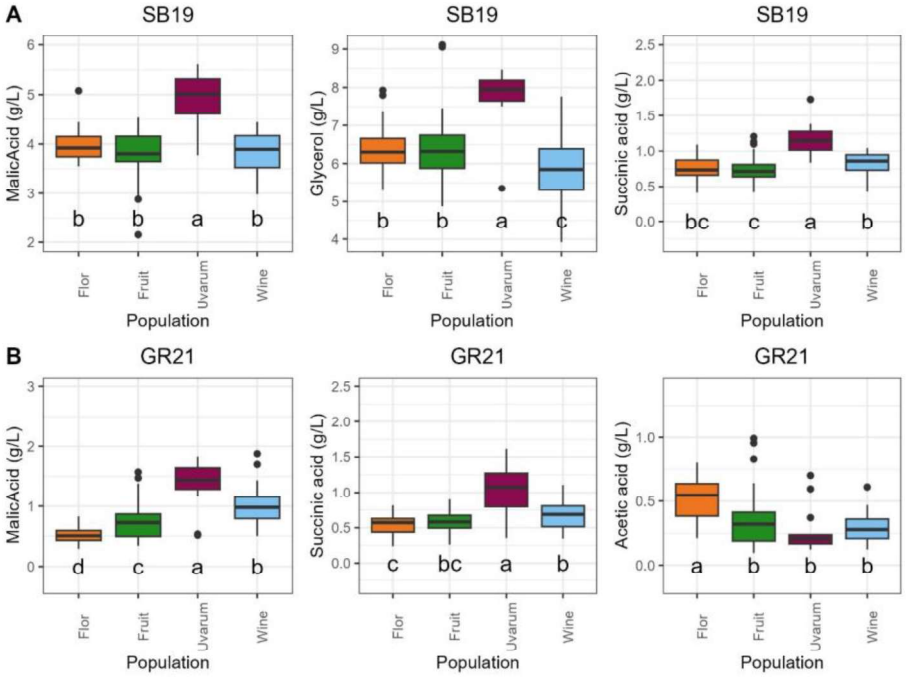
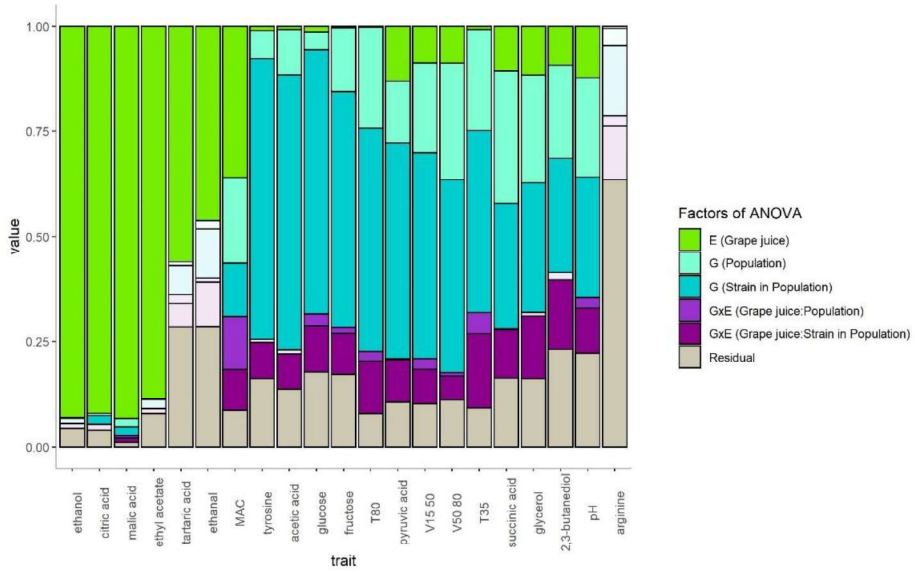



**Figure 4:** Main compounds involved in populations separation. (**A**) Absolute concentrations (g/L) of malic acid, glycerol, and succinic acid, in SB19 for the 4 populations. (**B**) Absolute concentrations (g/L) of malic acid, succinic acid, and acetic acid in GR21 for the four yeast populations. A Kruskal test has been applied to the compounds in both media. Different letters indicate a significant difference between the populations ($\alpha < 0.05$).




726
727 **Figure 5:** Bar graphs indicating the part of variance explained by the different factors of the ANOVA. The letters E and G
728 represent the environmental (grape juice) and the genetic factors, respectively. The nested ANOVA applied allows to evaluate
729 the effect of the population and the effect of strain within each population (strain). GxE represents the interaction between
730 genetic and environment and was decomposed in two nested factors grape juice: population and grape juice: strain in
731 population. Gray tons indicate non significative effect of the factors.

732

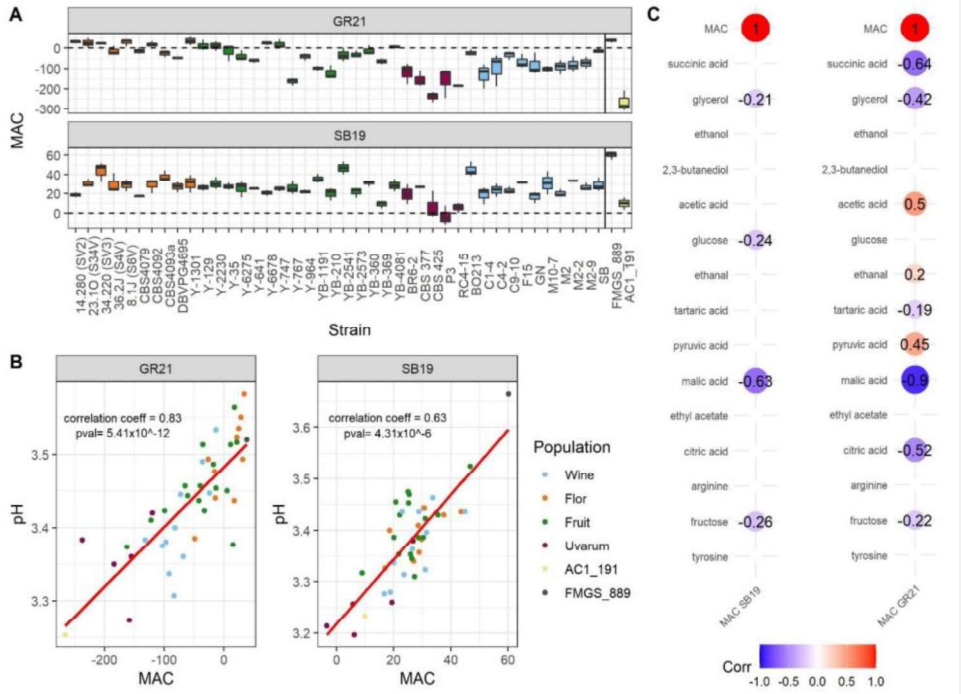



**Figure 6:** (**A**) MAC values measured for all strains in the two grape juices (3 replicates per strain). Strains were colored according to their population. The dashed line represents the 0% which corresponds to no malic acid consumed or produced. Positive values correspond to a consumption of malic acid and negative values to a production. (**B**) Correlation of MAC and pH for the two grape juices. The red line represents the linear regression line, correlation coefficients and p-values are written on the graphs. The dots represent all the tested strains (3 replicates per strains). They are colored according to their genetic group. (**C**) Correlation matrix between MAC and all other metabolites in the two grape juices. The value displayed corresponds to the correlation coefficient. Red values correspond to a negative correlation while blue values correspond to a positive correlation, only significative correlations ($\alpha < 0.05$) were indicated.

## Supplementary data

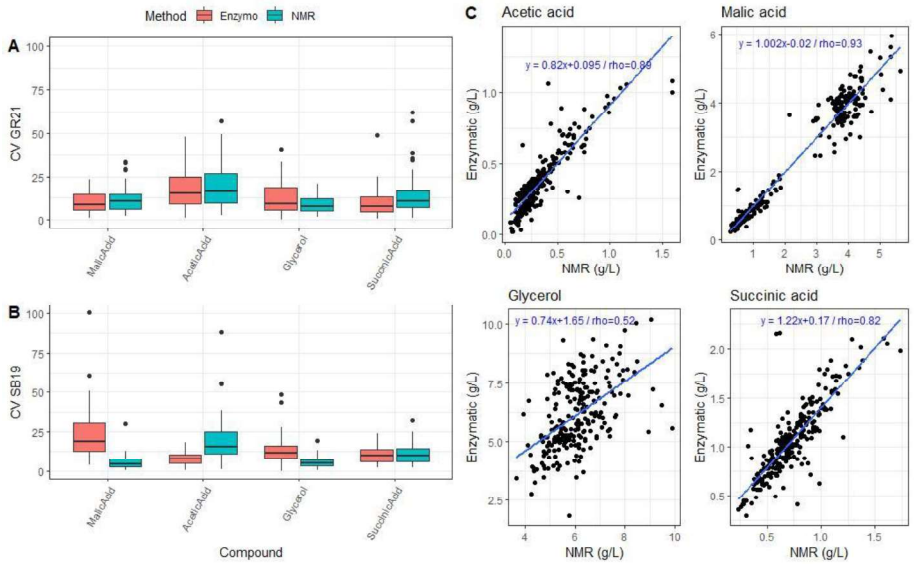



**Figure S1:** Correlation between enzymatic assay and NMR analysis for acetic acid, glycerol, malic acid, and succinic acid. The Spearman correlation coefficients are 0.89, 0.82, 0.93 and 0.52 for acetic acid, succinic acid, malic acid, and glycerol, respectively. The concentrations on both axes are expressed in g/L. There is an overestimation of the concentration of malic acid, succinic acid, and glycerol with the NMR method.

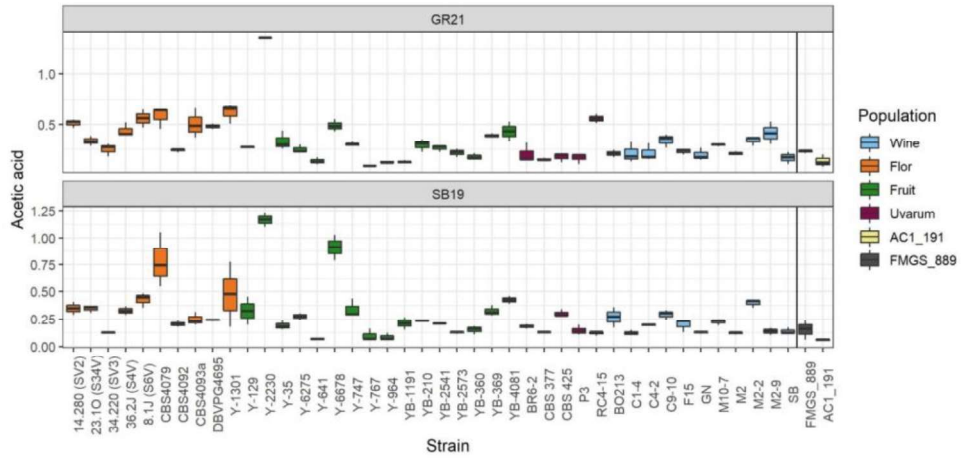



**Figure S2:** Absolute concentration in acetic acid (g/L) at the end of the fermentation for all strains in both grape juices. The boxplots are colored according to their population.

# Targeted $^{1}$-H-NMR wine analyses revealed specific metabolomic signatures of yeast populations belonging to the *Saccharomyces* genus.

Vion Charlotte[1,2], Le Mao Inès[2], Nadine Yeramian[3], Maïtena Muro[1,2], Margaux Bernard[1,2], Da Costa Gregory[2], Richard Tristan[2], Marullo Philippe[1,2]

[1]Biolaffort, Bordeaux, FRANCE

[2] UMR 1366 Œnologie, Université de Bordeaux, INRAE, Bordeaux INP, BSA, ISVV

[3]Microbiology Division, Department of Biotechnology and Food Science, Faculty of Science-University of Burgos, Spain

Corresponding author: Philippe Marullo

## Abstract

This study aimed to explore the non-volatile metabolomic variability of a large panel of strains (44) belonging to the *Saccharomyces cerevisiae* and *Saccharomyces uvarum* species in the context of the wine alcoholic fermentation. For the *S. cerevisiae* strains *flor*, fruit and wine strains isolated from different anthropic niches were compared. This phenotypic survey was achieved with a special focus on acidity management by using natural grape juices showing opposite level of acidity. A $^{1}$H NMR based metabolomics approach was developed for quantifying fifteen wine metabolites that showed important quantitative variability within the strains. Thanks to the robustness of the assay and the low amount of sample required, this tool is relevant for the analysis of the metabolomic profile of numerous wines. The *S. cerevisiae* and *S. uvarum* species displayed significant differences for malic, succinic, and pyruvic acids, as well as for glycerol and 2,3-butanediol production. As expected, *S. uvarum* showed weaker fermentation fitness but interesting acidifying properties. The three groups of *S. cerevisiae* strains showed different metabolic profiles mostly related to their production and consumption of organic acids. More specifically, *flor* yeast consumed more malic acid and produced more acetic acid than the other *S. cerevisiae* strains which was never reported before. These features might be linked to the ability of *flor* yeasts to shift their metabolism during wine oxidation.

## Key words

$^{1}$H-NMR, fermenting yeast, wine metabolome, malic acid, *flor* and wine yeast strain.

## 1. Introduction

Metabolic activities of microorganisms strongly impact the chemical composition of fermented products and modify their nutritional and organoleptic properties [1,2]. Thus, the quantification of chemical compounds in fermented goods is a critical step to understand the role of microbes and to control their development. In this context, wine alcoholic fermentation has been particularly well investigated, and many wine compounds were identified using GC-MS, LC-MS and NMR [3–6]. As widely reviewed [7–9], the yeast metabolism impacts the chemical composition of resulting wines with direct enological consequences. Although they broadly share the same metabolic pathways, yeast species and strains show a high variability in their metabolites production. Thus, yeast species involved in grape juice fermentation are characterized by specific metabolic signatures for primary metabolites [10,11], fermentative esters [12–14], and off-flavor compounds [15]. In mixed cultures, metabolic profiling is also useful for characterizing microbial interactions [16–18] that may impact wine complexity [19]. At the intra specific level, analytical chemistry methods are also decisive for driving yeast selection aiming to better control ethanol content [20], wine acidity [10,21], volatile thiols content [22] or fermentative esters production [13].

The natural genetic variability of the species *Saccharomyces cerevisiae* [23] constitutes an important source of metabolomic variability that has been deciphered for volatile [24,25] and non-volatile compounds [26,27] by applying quantitative genetic. These studies required the use of analytical methods such as HPLC [28], enzymatic assays [26], targeted GC-MS [24] allowing the quantification, for large sample sizes of wine metabolites belonging to different chemical family. Alternatively, untargeted metabolomics approaches may be applied for quantifying hundreds of metabolic features able to discriminate the biochemical signature of few *Saccharomyces cerevisiae* strains during the alcoholic fermentation [29][5]. However, the exact identification of such discriminating compounds remains a chemical challenge [30].

In this context, the exploration of extracellular metabolomic variability of yeast strains during the alcoholic fermentation for different classes of compounds constitute a challenging task. A performing and versatile analytical technique to quantify yeast metabolites is the proton Nuclear Magnetic Resonance ($^1$H NMR) spectroscopy [31,32]. The main $^1$H NMR spectroscopy assets are the simplicity of sample preparation and high reproducibility of the quantification. It is also possible to investigate a large range of metabolites belonging to different chemical families that are displayed on the same spectrum. This non-destructive technique also gives the opportunity to carry out several analyses on the same sample. NMR spectroscopy is quantitative since the signal intensity is directly proportional to the metabolite concentration and the number of nuclei in the molecule [33,34]. Despite its high potentiality to study fermented products, $^1$H NMR investigations are still scarce and face two major challenges [31]. First, the sensitivity of the technique is lower than mass spectrometry which explains that the latter is often preferred for non-targeted analyses. This issue can be addressed by applying a

high number of scans during the analysis, but this will extend the time required for the analysis. Secondly, peak overlaps from multiple metabolites pose major challenges. 2D NMR or supplementation with pure compounds of interest can be carried out to address this issue [35].

In the present study, we explored the metabolomic variability of different wine yeasts with a special focus on the management of wine acidity by malic acid. Indeed, the level of this organic acid tends to drastically decrease in grape juices due to climate change with important enological consequences [36]. As previously reported by several authors, strains and species of the *Saccharomyces* genus may modulate malic acid concentrations of wines [37–43]. Recently, pools of alleles belonging to *flor* or *wine* yeast populations were partially linked to the metabolic variability of malic acid [44] suggesting a possible effect of the yeast ecological origin. In this context, we reevaluated the natural variability of several *Saccharomyces* strains presenting different level of malic acid production or consumption during wine fermentation. In order to have a wide overview of possible metabolic trade-off related to malic acid metabolism, we developed new protocols aiming to quantify extracellular yeast metabolites using $^{1}$H-NMR based metabolomics, adapting methods already used for describing wine composition [45, 46].

## 2. Materials and Methods

### 2.1. Yeast strains used and culture methods

The forty-four strains of *S. cerevisiae* and *S. uvarum* isolated in different enological niches are listed in Table 1. *S. cerevisiae* strains were propagated on YPD 2 % (1 % peptone, 1 % yeast extract, 2 % glucose) at 30°C in both liquid and plate cultures (2 % agar). *S. uvarum* strains were propagated on YPD 6 % to avoid sporulation as described by [47]. Long term storage at -80°C was achieved by adding one volume of glycerol to YPD overnight cultures.

95 Table 1: Yeast strains used.

| Strain | Species | Ecological niche | Substrate/origin | Geographical area | Collection | Reference |
|---|---|---|---|---|---|---|
| FMGS1_889 | *S. cerevisiae* | Control | Extreme malic acid consuming strain | Miscelaneous (Breeding) | UMR Oenology (ISVV) | [21] |
| FMGS3_191 | *S. cerevisiae* | Control | Extreme malic acid producing strain | Miscelaneous (Breeding) | UMR Oenology (ISVV) | [70] |
| 14.280 (SV2) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 23.10 (S34V) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 34.220 (SV3) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 36.2J (S4V) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| 8.1J (S6V) | *S. cerevisiae* | Flor | Velum | Dijon, France | Prof. Hervé Alexandre | [71] |
| DBVPG4695 | *S. cerevisiae* | Flor | Vino Santo | Lungarotti winery, Italy | 1002 genomes | [23] |
| CBS4079 | *S. cerevisiae* | Flor | Velum | Spain | 1002 genomes | [23] |
| CBS4092 | *S. cerevisiae* | Flor | Velum | Spain | 1002 genomes | [23] |
| CBS4093a | *S. cerevisiae* | Flor | Velum | Spain | 1002 genomes | [23] |
| Y-1301 | *S. cerevisiae* | Flor | Wine | Unknown | NRRL collection | |
| YB-210 | *S. cerevisiae* | Fruit | Spoiled banana | Costa Rica | NRRL collection | |
| Y-747 | *S. cerevisiae* | Fruit | Cider | Illinois, USA | NRRL collection | |
| YB-1191 | *S. cerevisiae* | Fruit | Citrus juice | Louisiana, USA | NRRL collection | |
| Y-6678 | *S. cerevisiae* | Fruit | Olives | Spain | NRRL collection | |
| Y-2230 | *S. cerevisiae* | Fruit | Fruit juice | The Netherlands | NRRL collection | |
| YB-360 | *S. cerevisiae* | Fruit | Applesauce | Unknown | NRRL collection | |
| YB-2541 | *S. cerevisiae* | Fruit | Benzolated cider at 18C | Unknown | NRRL collection | |
| Y-641 | *S. cerevisiae* | Fruit | Cider | Unknown | NRRL collection | |
| Y-35 | *S. cerevisiae* | Fruit | Fruit (Ilex aquifolium) | Unknown | NRRL collection | |
| Y-6275 | *S. cerevisiae* | Fruit | Orange concentrate | Unknown | NRRL collection | |
| Y-129 | *S. cerevisiae* | Fruit | Orange juice | Unknown | NRRL collection | |
| YB-4081 | *S. cerevisiae* | Fruit | Ripe goyave | Unknown | NRRL collection | |
| YB-2573 | *S. cerevisiae* | Fruit | Sauerkraut | Unknown | NRRL collection | |
| YB-369 | *S. cerevisiae* | Fruit | Sauerkraut | Unknown | NRRL collection | |
| Y-964 | *S. cerevisiae* | Fruit | Sour figs | Unknown | NRRL collection | |
| Y-767 | *S. cerevisiae* | Fruit | Tomato product (B-117) | Unknown | NRRL collection | |
| RC4-15 | *S. cerevisiae* | Uvarum | Wine | Alsace, France | UMR Oenology (ISVV) | [49] |
| BR6-2 | *S. uvarum* | Uvarum | Cider | Britany/Normandy, France | UMR Oenology (ISVV) | [49] |
| CBS 377 | *S. uvarum* | Uvarum | Fruit juice | Germany | UMR Oenology (ISVV) | [49] |
| P3 | *S. uvarum* | Uvarum | Wine | Sancerre, France | UMR Oenology (ISVV) | [49] |
| CBS 425 | *S. uvarum* | Uvarum | Cider | Switzerland | UMR Oenology (ISVV) | [49] |
| GN | *S. cerevisiae* | Wine | meiotic spore clone from Zymaflore VL1 | Bordeaux, France | UMR Oenology (ISVV) | [43] |
| C1-4 | *S. cerevisiae* | Wine | Wine | Cordoba, Spain | UMR Oenology (ISVV) | [40] |

| | | | | | | |
|---|---|---|---|---|---|---|
| C4-2 | *S. cerevisiae* | Wine | Wine | Cordoba, Spain | UMR Oenology (ISVV) | [40] |
| C9-10 | *S. cerevisiae* | Wine | Wine | Cordoba, Spain | UMR Oenology (ISVV) | [40] |
| SB | *S. cerevisiae* | Wine | meiotic spore clone from Actiflore BO213 | France | UMR Oenology (ISVV) | [43] |
| F15msp | *S. cerevisiae* | Wine | meiotic spore clone from Zymaflore F15 | France | UMR Oenology (ISVV) | [72] |
| M10-7 | *S. cerevisiae* | Wine | Wine | Madrid, Spain | UMR Oenology (ISVV) | [40] |
| M2-2 | *S. cerevisiae* | Wine | Wine | Madrid, Spain | UMR Oenology (ISVV) | [40] |
| M2-9 | *S. cerevisiae* | Wine | Wine | Madrid, Spain | UMR Oenology (ISVV) | [40] |
| M2msp | *S. cerevisiae* | Wine | meiotic spore clone from Enoferm M2 (Lallemand) | Unknown | UMR Oenology (ISVV) | [72] |
| BO213 | *S. cerevisiae* | Wine | Wine starter | France | Laffort | [26] |

### *2.1.1. Grape Juices*

Two grape juices, Sauvignon blanc 2019 (SB19) and Grenache 2021 (GR21) were collected in Bordeaux and Montpellier areas, respectively and were stored at -20°C. The GR21 is a red grape juice representing warm climate matrix with a very low malic acid content. The SB19 is a white grape juice that was supplemented with L-malic acid in order to artificially increase its acidity. The final composition of these grape juices in fermenting sugar, malic acid content and pH is listed in Table 2.

Table 2: Composition of the different grape juices used in the experiments.

| **Composition** | **SB19** [a] | **GR21** |
|---|---|---|
| L-malic acid (g/L) | 5.31 | 0.52 |
| pH | 3.20 | 3.52 |
| Fermentable sugars (g/L) | 202.4 | 240.0 |
| Assimilable nitrogen (mg N/L) | 124 | 106 |
| Total $SO_2$ (mg/L) | 53 | 32 |

[a] *The SB19 has been supplemented to increase its malic acid content and decrease the pH. Its original composition was a malic acid concentration at 3.73 g/L and a pH of 3.38.*

### *2.1.2. Alcoholic fermentation monitoring*

Small-volume alcoholic fermentations were implemented in screwed vials fermentations according to the general procedure described in [48]. Rapidly, 20 mL-screwed vials (Thermo Fisher Scientific, Bordeaux, France) were filled with 12 mL of grape juice and were tightly closed with screw cap-magnetic (Agilent Technologies, hdsp cap 18 mm PTFE/il 100 pk, Les Ulis, France) perforated

with hypodermic needles (G26-0.45 x 13 mm, Terumo, Shibuya, Tokyo, Japan) for allowing $CO_2$ release. Vessel was inoculated by $2.10^6$ viable cell.mL$^{-1}$ precultured in liquid media 50 % filtered must, 50 % sterile $H_2O$ for 24h. Cellular concentration and viability was estimated by flow cytometry using a CytoFlex (Beckman Coulter, Villepinte, France). Fermentation took place at 24°C in shaken vials by using an orbital shaker (SSL1, Stuart, Vernon Hills, IL, USA) at 175 rpm. Fermentation kinetics were estimated by monitoring manually (1-2 times per day) the weight loss caused by $CO_2$ release using a precision balance with automatic weight recording (LabX system, Mettler Toledo, Viroflay, France). The amount of $CO_2$ released according to time was modeled by local polynomial regression fit [48]. This model allows the estimation of the time necessary to reach the maximum $CO_2$ produced, the lag phase, the speed between 15 % and 50 % of the fermentation (V15_50), the speed between 50 % and 80 % (V50_80), the time to reach 80 % of $CO_2$ produced (T80) and the maximal theoretic $CO_2$ produced ($CO_2$max). Final pH was monitored using Five Easy Plus pH-meter (Mettler Toledo, Viroflay France) with a micro probe LE422 (Mettler Toledo, Viroflay France).

### 2.2. Microsatellite genotyping

The genomic DNA of the 39 *S. cerevisiae* strains was quickly extracted in 96-wells microplate format using a customized LiAc-SDS protocol [49]. Fourteen polymorphic microsatellite loci (*SCAAT3*, *C3*, *C5*, *SCYOR267C*, *C8*, *C11*, *SCAAT2*, *YKL172*, *C9*, *C4*, *SCAAT5*, *SCAAT1*, *C6*, *SCAAT5*, *SCAAT1*, *C6*, *YPL009*) were used for estimating the genetic relationships within those strains using PCR conditions previously described [50]. Two multiplex PCRs allowing genotyping of seven loci were carried out in a final volume of 12.5 µL containing 6.25 µL of Qiagen Multiplex PCR master mix and 1 µL of DNA template. 1.94 µL of each mix was added in the mixture using the concentrations indicated. Both reactions were run with the following program: initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 57°C for 2 min, 72°C for 1 min, and a final extension at 60°C for 30 min. The size of PCR products was analyzed by the MWG company (Ebersberg, Germany) using 0.2 µL of 600 LIZ (GeneScan) as a standard marker. Chromatograms were analyzed with the GeneMarker (V2.4.0, Demo) program.

### 2.3. Metabolites analysis by $^1$H NMR

#### *2.3.1. Wine samples preparation*

Samples were collected during the alcoholic fermentation. Sixty µL of centrifuged wine sample was diluted ten times in phosphate buffer prepared in $D_2O$ (pD 4.2, 0.1 M final). Sixty µL of solution of calcium formate (FCa) 1 mM and trimethylsilylpropanoic acid (TSP) 1.25 mM in $D_2O$ was added to the preparation. The FCa is used as an internal standard for calculating the concentration of metabolites and the TSP to set the spectrum at 0 ppm.

#### *2.3.2. NMR spectra acquisition*

Spectra were recorded on a 600 MHz Avance III NMR spectrometer (Bruker, Wissembourg, France) operating at 600.25 MHz, equipped with a TXI 5 mm probe with z gradient coils. The measurement was performed at 293 K using TopSpin 4.0.8 software (Bruker, Wissembourg, France). A 1D-NOESY pulse sequence (*noesygppr1d*) was used with low power presaturation at the water frequency during relaxation delay and mixing time and spoil gradient. Relevant parameters were: spectral size: 64k; number of scans: 64; spectral width: 18 ppm; acquisition time: 3.64 s; relaxation delay: recovery delay: 5 s; mixing time: 100 ms. The 90° pulse calibration was carried out for each sample automatically, and the shimming was set manually in *gs* mode for each spectrum in order to obtain the finest possible line width (lower than 1 Hz). The FIDs were multiplied by an exponential function corresponding to a 0.3 Hz line-broadening factor prior to the Fourier transformation. Manual phase followed by automatic baseline corrections were applied to the resulting spectrum, which was aligned to zero using the TSP signal.

Wine metabolite identification was performed using databases, literature data [51], and addition of pure standards when faced with uncertainties. Additionally, some 2D experiments were used (TOCSY and HSQC) to insure the identification. The FCa signal at 8.28 ppm was used as internal standard for calculating the concentration of the identified compounds. Peak deconvolutions were performed by the global spectral deconvolution method (GSD) [52], using the simple mixture analysis (SMA) plugin of MestReNova 12.0 software (Mestrelab Research, Santiago de Compostela, Spain). Quantification was achieved according to the formula of Goldelmann *et al.* [62].

## 2.4. Statistical analyses

All the statistical and graphical analyses were carried out using R software [53] and plots were generated using the base or *ggplot2* packages.

### *2.4.1. Phylogenic analysis*

The microsatellite dataset was manipulated using the *adegenet* package implemented in R and the genetic distance within the strains was estimated using the Bruvo's distance using the *poppr* package [73]. The phylogram was built by Neighbor Joining (*ape* package) [74]. In order to evaluate a possible genetic structuration between *fruit*, *wine* and *flor* yeasts, a Discriminant Analysis of the seventh Principal Components was applied using the *adegenet* package omitting the two control strains (FMGS_889 and AC1_191). The pairwise *Fst* between populations was estimated using the *genet.dist* function.

### *2.4.2. Multivariate analyses*

Spearman correlation tests between traits were computed using the *corr* function (*psych* package) corrected for multiple tests using the Benjamini-Hochberg method ($\alpha = 0.001$). Results were displayed with the *corrplot* function (*corrplot* package). The multivariate phenotypic variability of yeast strains was visualized by a Principal Component Analysis (PCA) using the *ade4* package. In order to

highlight the most discriminating variables of PCA, Discriminant Analysis of Principal Component (DAPC) (*adegenet* package) was applied by selecting principal components allowing to capture up to 60% of the cumulated inertia.

#### *2.4.3. Analysis of variance*

Analyses of Variance (ANOVA) were carried out using the *car* package. The phenotypic values measured in the grape juices SB19 and GR21 were analyzed using the linear model (LM1) in order to estimate the effect of the following factors: media, population, strain in population as well as the first order interaction of all the factors according to the formula (1).

(1) $y_{ijk} = media_i + population_j + strain(population)_{jk} + inter^2_{ijk} + \varepsilon_{ijkl}$

where *y* is value of all the variables for a media *i* (*i*=1,2) in which *j* (*j*=1,2,3,4) groups of yeast strains fermented. Each population is composed of *k* strains. The factor strain is nested in the factor population and *k* varies between 1:5 and 1:16 according to the number of strains per population. The term $inter^2_{ijk}$ represents the first order interaction of each factor and $\varepsilon_{ijkl}$ the residual. The analysis of variance of model LM1 allows the estimation of the primary effect of the media, population, and strain in population on several quantitative variables, as well as their primary interaction effect. The normal distribution of residues as well as the homoscedasticity of variances were tested by Shapiro test and Levene test (*car package*), respectively. When necessary, non-parametric comparison of samples were carried out using the Wilcoxon-Mann-Whitney or Kruskal test with corrected p values (Benjamini-Hochberg method, $\alpha = 0.05$).

## 3. Results

### 3.1. Genetic characterization of a panel of *Saccharomyces* strains

The forty-two strains used were classified according to their origin and were denominated thereafter "*flor*" (n=10), "wine" (n=11), "fruit" (n=16), and "uvarum" (n=5) (Table 1). *Flor* and wine yeast populations (*S. cerevisiae*) share the same ecological niches but have been clearly separated by molecular phylogenetic studies [54,55]. As mentioned in the introduction, several genes related to central metabolism and proton homeostasis have been linked to the ability to consume, or not, malic acid during the alcoholic fermentation [44]. For most of them, alleles of the *flor* origin were related to a stronger consumption of malic acid suggesting a possible metabolic adaptation of this population to malic acid. However, phenotypic differences of malic acid consumption between *flor* and wine yeasts were never reported to our knowledge. To enlarge the genetic diversity of this metabolic survey, several *S. cerevisiae* strains (16) related to the fermentation of fruits (agrums, apple) and acidic substrates (sauerkraut or tomato juice) were included in the panel. Finally, five *S. uvarum* strains presenting different genetic origin [42] were integrated to this panel since this psychrophilic species of the *Saccharomyces* genus is a strong producer of malic acid [56]. In addition, two *S. cerevisiae* strains AC1_191 and FMGS_889 were included to this panel as "control" for a total of 44 strains. These last

two strains have been recently selected for their extreme ability to consume or produce malic acid [21,57] and have a mixed inheritance between *flor* and wine origin. The Bruvo's genetic distance between the 37 *Saccharomyces cerevisiae* strains used was computed by using 14 microsatellites markers as previously described [26]. The genetic distances range is between 0.11 and 0.96, indicating that all the strains are unique as illustrated by the genetic tree shown in Figure 1A with some confusions between *flor* and wine populations. A discriminant analysis of the ten first Principal Components (60 % of the cumulative variance) allows the separation of the three populations (Figure 1B) with a probability of assignation of 0.90, 1.00 and 0.93 for *flor*, wine, and fruit populations respectively. The pairwise Fst between *flor* and wine strains was 0.169 indicating a clear separation between the two populations. The five strains of *S. uvarum* are also unique and represent the variability of holoarctic *S. uvarum* as reported in a former study [42].

## 3.2. Assessment of wine metabolites variability of fermenting yeasts by targeted $^{1}$H-NMR analysis.

### *3.2.1. $^{1}$H-NMR analysis of yeast*

A targeted $^{1}$H-NMR metabolomics approach was applied for quantifying metabolites of fermenting yeasts belonging to the *Saccharomyces* genus. The main wine metabolites produced and consumed by yeasts during the alcoholic fermentation were quantified from sixty microliters of wine stored at -80°C. Samples were simply thawed and diluted in an appropriate buffer before analysis as detailed in methods. The typical $^{1}$H-NMR spectrum after water suppression is presented in Figure 2. The signals at 0.00 ppm and 8.28 ppm correspond to TSP and FCa respectively; other signals correspond to wine constituents. The $^{1}$H-NMR spectra were dominated by ethanol, and glycerol, followed by organic acids. Even if different amino acids were observed in juice, due to their consumption by yeast [58], only tyrosine and arginine were quantified after the alcoholic fermentation. The Table 3 shows the chemical shifts and the coupling constants used for identification and quantification of 15 metabolites including six organic acids (acetic acid, citric acid, malic acid, pyruvic acid, succinic acid, and tartaric acid), three alcohols (ethanol, 2,3-butanediol, glycerol), two reducing sugars (fructose and glucose), one ester (ethyl acetate), one aldehyde (ethanal), and two amino acids (tyrosine and arginine). The concentration range and the average coefficient of variation (CV) of each metabolite measured in one red (GR21) and one white grape juice (SB19) is also indicated, as well as the average concentration in both grape juices.

**Table 3:** Typical chemical shifts and coupling constants used for compounds identification after fermentation. The signals chosen for quantification are in bold. Variation coefficients and average concentration are displayed for all compounds in both media.

| Compound | δ1 H (Multiplicity, J in Hz, Assignment) | Average CV SB19 | Relative Average concentration SB19 (g/L) | Average CV GR21 | Relative Average concentration GR21 (g/L) |
|---|---|---|---|---|---|
| acetic acid | **2.08** (s, CH3) | 21.06 | 0.2786 | 20.27 | 0.3166 |
| arginine | **1.68** (m, CH2) | 8.67 | 0.0031 | 43.09 | 0.0074 |
| 2,3-butanediol | **1.13** (d, 6.2, 2CH3), | 15.84 | 0.4027 | 17.77 | 0.5089 |
| citric acid | **2.69** (d, 15.6, CH2), | 7.49 | 0.4265 | 49.74 | 0.0953 |
| ethanal | **2.23** (d, 3.0, CH3), 9.79 (q, 2.85, CH) | 60.44 | 0.0023 | 34.15 | 0.0141 |
| ethanol | **1.17** (t, 7.2, CH3), 3.65 (q, CH2) | 7.89 | 59.6569 | 4.82 | 104.2476 |
| ethyl acetate | 1.26 (t, 7.2, CH3), **2.03** (s, CH3), 4.12 (q, CH2) | 8.48 | 0.7078 | 5.97 | 1.1418 |
| fructose | 3.97 (dd, 10.1, 3.5, CH), **4.01** (m, CH), 4.09 (dd, 12.8, 1.0, CH) | 37.38 | 1.3929 | 49.99 | 1.8364 |
| glucose | **4.65** (d, 7.9, CH), **5.23** (d, 3.6, CH) | 25.47 | 0.2182 | 32.71 | 0.1882 |
| glycerol | **3.55** (dd, 11.8 and 6.5, 2CH2), 3.77 (m, CH) | 5.50 | 10.1624 | 9.01 | 9.0169 |
| malic acid | 2.59 (dd, 16.3 and 7.0, CH), **2.81** (dd, 16.3 and 4.5, CH), 4.36 (dd, CH) | 5.44 | 6.9769 | 12.28 | 1.2349 |
| pyruvic acid | **2.38** (s, CH3) | 18.14 | 0.0542 | 21.83 | 0.0341 |
| succinic acid | **2.58** (s, 2CH2) | 10.81 | 1.8568 | 15.24 | 1.4810 |
| tartaric acid | **4.39** (s, 2CH) | 40.52 | 0.5950 | 15.31 | 1.1052 |
| tyrosine | **6.86** (m, 2CH), 7.17 (m, 2CH) | 16.34 | 0.0470 | 15.05 | 0.0517 |

In both grape juices, most compounds had a CV lower than 30 % and in the SB19, six compounds had a CV under 10 % which shows the good repeatability and reliability of the analysis. Some compounds (residual sugars, ethanal, arginine and tartaric acid) were not accurately quantified due to their low concentrations and their instability to sample concentration (see discussion).

The relative concentrations presented in the table 3 allows the statistical comparisons between samples. For some key compounds (acetic acid, malic acid, succinic acid, and glycerol), we quantified their absolute concentrations by applying correcting factors calculated using the standard addition method and applied using the formula of Goldelman et al. [62] (table 4). This absolute quantification was compared to enzymatic assay methods commonly used in enology [43]. As expected, strong correlations between enzymatic and $^{1}$H-NMR quantifications were found for all the metabolites (Spearman correlation analysis). The average CV of both methods were similar in the two grape juices, demonstrating that $^{1}$H-NMR assay was also very reproducible for such metabolites (Figure S1 panels A and B). For acetic acid, malic acid, and succinic acid, the high correlations coefficient (rho >

0.82) indicated a good level of agreement between the two methods. Glycerol concentrations were also correlated despite a weaker Spearman's coefficient (rho = 0.52, p-value = $2.79x10^{-21}$). Linear regressions suggested that quantification by $^{1}$H-NMR of acetic acid and glycerol were slightly overestimated respect to enzymatic assays (slopes of 0.82 and 0.74, respectively). In contrast succinic acid concentration was slightly overestimated (slope of 1.22). However, both methods resulted in very similar quantifications and were in agreement with enological values expected.

**Table 4:** Correction factors and absolute average concentration measured for acetic, malic, succinic acids and glycerol in red and white wines.

| Compound | Correction factors | | Absolute average concentration SB19 (g/L) | Absolute average concentration GR21 (g/L) |
|---|---|---|---|---|
| | SB19 | GR21 | | |
| acetic acid | 0.94 | 1.17 | 0.26 | 0.37 |
| glycerol | 0.63 | 0.64 | 6.40 | 5.77 |
| succinic acid | 0.44 | 0.44 | 0.82 | 0.65 |
| malic acid | 0.57 | 0.68 | 3.97 | 0.84 |

### *3.2.2. Multivariate analysis of wine metabolites*

Fermentations were monitored daily by following weight loss, and six kinetic parameters (lag phase, V15_50, V50_80, T35, T80 and $CO_2$max) were extracted as previously reported [48]. The fifteen end-point metabolites were quantified for a panel of 44 strains of different origins (Table 1) in two grape juices (SB19 an GR21) fermented in triplicate. A Principal Component Analysis (PCA) was applied to average values in order to capture the overall variability of the 15 wine metabolites quantified by $^{1}$H-NMR (Figure 3). The final concentrations of wine metabolites are mostly structured by the grape juice nature and are clearly separated by the first component (40.3 % of inertia). Indeed, the SB19 was enriched in malic acid and citric acid compared to the GR21. In contrast, GR21 displayed a higher sugar content resulting in a higher production of ethanol, ethyl-acetate and ethanal. Beside this grape juice effect, *S. uvarum* and *S. cerevisiae* species are partially separated by the second component which captured 16.9 % of the total inertia. This axis is mostly correlated with glycerol and succinic acid content which are overproduced by *S. uvarum* strains. In addition, the presence of reducing sugars at the end of the alcoholic fermentation was detected for some S. *uvarum* and *fruit* strains which is also strongly related to the axis 2. In contrast, the quantification of the 15 metabolites did not allow a clear separation of the three *S. cerevisiae* populations in both matrices.

In order to identify metabolic signatures able to better discriminate yeast populations, the two grape juice datasets were analyzed separately using a Discriminant Analysis of Principal Component (DAPC). In the SB19, malic acid is the most discriminating compound followed by glycerol and succinic acid. The cumulated variability explained by these three metabolites on main linear discriminant axis is 98.3 %. As shown in Figure 4A, wines fermented by the *S. uvarum* species are

significantly more concentrated in malic acid, glycerol, and succinic acid (Kruskal test, alpha < 0.05) than *S. cerevisiae*. These two species were also well discriminated in GR21 grape juice by three compounds: malic acid, succinic acid, and acetic acid which represent 64 %, 31 %, and 2 % of the discriminating inertia, respectively (Figure 4B). Beside this major species effect, significant differences of minor magnitudes were also observed within *S. cerevisiae* populations with a special emphasis for wine and flor populations that were always significantly different. In GR21 grape juice, the four populations were better separated than in the SB19 grape juice. Indeed, in GR21 each population was statistically different for malic acid content (Kruskal-Wallis' test, p-value < 0.05) while in SB19, only *S. uvarum* produced significantly more malic acid than the three other groups. This result could be linked to the initial characteristics of GR21 that contained a very low amount of malic acid.

### *3.2.3. Contributions of genetic and environmental effect on the phenotypic variability of Saccharomyces strains*

To deeply investigate factors influencing kinetic and metabolic traits, a nested analysis of variance was applied, aiming to estimate the impact of environment (grape juice) and genetic factors and their possible interactions. The genetic contribution effect was decomposed in population effect and strain within population effect as detailed by the linear model 1 (see methods). The contribution of each factor on 21 quantitative traits is summarized in Figure 5. The variability of ethanol, citric acid, malic acid, ethyl acetate, tartaric acid, and ethanal was mainly due to the grape juice effect. Indeed, ethanol and ethanal are directly linked to the initial sugars concentration that differs between red and white grape juices (SB19 = 202.4 g/L, GR21 = 240 g/L). Ethyl acetate is a carboxylate ester which is a secondary metabolite produced during alcoholic fermentation, derived from ethanol [59]. Even though its production can vary according to the strain, the very different ethanol content between the two wines led to significantly different amount of ethyl acetate. Tartaric acid is almost two times more concentrated in GR21 than SB19 (1.10 g/L vs 0.59 g/L) and the average amount of citric acid is about 4 times more important in SB19 than in GR21 (0.426 g/L vs 0.095 g/L respectively). Finally, malic acid varied greatly with the environment as the two grape juices were selected for their extreme acidities.

To better estimate the genetic contribution on malic acid content, we used the variable *MAC* (Malic Acid Consumed) that represents the ratio of malic acid consumed expressed in percentage (Figure 6A). *MAC* variability was evenly influenced by G (33 % var), E (36 % var) and GxE interactions (22 % var). The *MAC* values of the two control strains AC1_191 [57] and FMGS_889 [21] highlighted their great impact on acidity management since they produced and consumed significantly more than any other *S. cerevisiae* strains of this study. Quantitatively, the differences of absolute malic acid concentrations between wines made by these extreme strains were 1.71 and 2.95

g/L for GR21 and SB19, respectively. Intuitively, the *MAC* is strongly correlated to the final pH (Figure 6B) since wine pH depends on the concentration of grape juice organic acids and especially the amount of malic acid. The Figure 6C shows the correlation of *MAC* with other metabolites in both grape juices. Overall, malic producer strains such as *S. uvarum* and AC1_191 produced more succinic acid than other.

For other metabolites (acetic acid, 2,3-butanediol, fructose, glycerol, pyruvic acid, succinic acid, and tyrosine) as well as kinetics parameters (V15_50, V50_80, T35 and T80), the genetic effect is greater than the grape juice's one. The decomposition of genetic effect in population and strain contribution indicated that within the same population, the metabolic variability of strains is generally stronger than the metabolic variability observed between groups (Figure 5). This is well illustrated by acetic acid variability that is more explained by the strain within populations (65 %) than by population effect itself (11 %). This can be explained for instance by some *flor* (CBS4079) and fruit (Y-2230, Y-6678) strains that overproduced acetic acid compared to other strains from their own populations (Figure S2). Finally, arginine was the unique metabolite that was neither impacted by grape juice or yeast strain, likely due to its low concentration level at the end of the fermentation.

# 4. Discussion

## 4.1. $^{1}$H NMR as an effective tool for analyzing wine microorganism metabolites.

Alcoholic beverages are complex matrices composed of several hundred volatile and non-volatile molecules that participate to the overall quality of the product. Therefore, analytical chemistry efforts aiming to characterize such complex matrices must be done for understanding and quantifying the role of microorganisms that participate to their elaboration. From a methodological viewpoint, trade-offs exist between the number of biological samples to analyze, the number of compounds assayed, the volume of sample required for their quantification, and the cost of the analysis. Generally, most of the microbiology studies lay emphasis on the number of biological sample and conditions analyzed in order to better understand the complex interactions existing between microbiological diversity and environmental conditions. This wide phenotypic characterization requires a reduction of fermentation volumes that allows a more efficient parallelization of culture conditions [43]. However, this strategy generally reduced the number of metabolites investigated by reducing the available sample volume. Thus, it is necessary to use high throughput methods such as enzymatic assay [43] or HPLC analyses [28]. Enzymatic assays are cost effective and can be easily robotized, however they require a specific assay per compounds with a quite limited number of metabolite available. HPLC analyses are more expensive and require the use of specific methods for quantifying organic acid and sugars [60], or nitrogen compounds [61].

NMR metabolomics is particularly amenable to detect compounds that are less tractable by liquid chromatography such as sugars, organic acids, alcohols polyols, and other highly polar compounds

[33]. Those classes of compounds are well represented in wine reinforcing the interest of this technique. In this study, we used an analytical method able to quantify from a small volume (less than 100 µL) using quantitative $^{1}$H-NMR spectroscopy. The advantages of NMR are the simplicity of sample preparation, measurement rapidity and the possibility to detect compounds belonging to different chemical families on one spectrum, in a single experiment. Previous studies were focused on the characterization of commercial wines and analysis of the cultivar, geographic origin, or vintage [45,46,64]. Our study focused on the characterization of wines fermented in the laboratory with different yeasts strains. The $^{1}$H-NMR method allowed the identification and quantification of 15 wine soluble metabolites of 264 biological samples requiring only 60 µL of wine for each experiment. This original approach allows studying the origin of many strains and their impact on various wine metabolites in different enological matrices.

The NMR analytical method was cross assessed by comparing the quantifications obtained by enzymatic assays for three organic acids (malate, succinate, and acetate) and for glycerol. Data obtained were accurately correlated for these metabolites and both methods had low coefficients of variation. By applying correction factors experimentally determined. For all the four compounds assayed the correlation between methods was very satisfactory and the slope of the linear models close to 1 +/- 20 % (Figure S1). For other compounds the absolute quantification was not determined but the relative concentrations showed in the table 3 did not impair the relative comparison of the strains which is the scope of this study. The quantification of glycerol seems more problematic even if the coefficient of variation of repetition is quite low (5.5% in SB19 and 9.01% in GR21). As shown Figure S1, the $^{1}$H-NMR method over-estimated glycerol content in comparison to enzymatic assays. The relative discrepancy between the two methods could be due to several factors relating to both methods. Concerning NMR, the signal of glycerol ($\delta_H$ = 3.55 ppm) is in a very dense region of the spectrum. This region is dominated by ethanol signal at 3.65 ppm. An overlap might have led to uncertainty of measurement.

In the present study, ethanol was quantified on all spectra as it is a compound of major interest in phenotypic analysis. However, the high intensity of the ethanol signal masks the signal of other compounds such as acids present in lower quantity (γ-aminobutyric acid, galacturonic acid, glucuronic acid, sorbic acid…). Therefore, the sensitivity of this method could be optimized in order to identify a higher number of compounds by suppressing the ethanol signal which is dominant after water using a selective pulse. Another solution would consist of introducing a lyophilization step to suppress the water and ethanol signals. This would improve the signal-to-noise ratio and better observe non-volatile compounds close to the signals of water and ethanol. These extra steps would increase the time of preparation and analysis but would allow the identification and accurate quantification of a higher number of other compounds.

Interestingly, the method applied is also useful for quantifying compounds that are not easily quantified by classical analyses such as ethyl acetate and ethanal. Ethyl acetate can be of interest as it

is the major ester in several distilled spirits such as whiskey, rum and cachaça [63,64]. Ethanal, is found in wines in various concentrations and can be formed by yeasts or be the indication of a contamination by acetic acid bacteria, lactic acid bacteria or be the result of auto-oxidation of ethanol and phenolic compounds [65]. The monitoring of its concentration in wines during fermentation and aging could be useful as its presence in small concentration gives a pleasant fruity aroma but at high concentrations it gives a pungent irritating odor. Finally, $^{1}$H-NMR analytical method is also applicable to grape juices and ongoing fermentations. The present study focused on the analysis of wines, but samples taken earlier during alcoholic fermentation could allow the assay of nitrogen compounds which are more abundant in the earliest stages of fermentation. On some grape juices' spectra, amino acids such as leucine, isoleucine, valine, threonine, arginine, proline, and choline were clearly observed (personal communication) in the 0.9 – 3.3 ppm region of the spectrum. Due to their low concentration and consumption by yeast, these compounds were not quantified in fermented wines, except for arginine which is initially present in high concentration in grape juices and is lately consumed by yeasts [58].

### 4.2. Wine metabolome is partially impacted by the population origin of the fermenting strains.

The metabolomic characterization of two grape juices fermented by a large panel of yeast strains was achieved. Those strains provided by microorganism collections are associated with acidic fermented matrices such as wine, cider, fruit juices, and sauerkraut. As confirmed by microsatellite analysis, the four groups of strains collected in different ecological or human associated niches are discriminated by population genetics tools. The three *S. cerevisiae* groups belong to distinct populations (*flor*, *wine*, and *fruit*) the last group being constituted by *S. uvarum* strains isolated in cider and wine environments. This last species has been reported to have acidifying properties and to produce large quantities of malic acid [42].

In the context of this study, the impact of the yeast strain on wine metabolome was investigated in divergent matrices (one red and wine white grape juice) showing opposite level of malic acid content (0.52 g/L vs 5.31 g/L). Initially, the GR21 had a low content in malic acid compared to SB19 which enhanced the ability of some strains to produce more malic acid. As malic acid has a buffer effect, an addition of malic acid to the medium reinforced this buffer effect. On the contrary, in SB19 most strains consumed malic acid, limiting the buffer effect. The correlation between MAC and pH is strong in both grape juices (Figure 6B) as malic acid is the organic acid influencing the most the pH. Tartaric acid is also a strong influencer of the acidity, but it is not modulated by yeasts during alcoholic fermentation. Indeed, the figure 5 shows that this organic acid is only impacted by the grape juice and there is no significant effect of the strains or populations.

The main insight of this study is that metabolome variations are partially structured by the origin of the strain. First, the phenotypic discrepancies between the two main species involved in the

alcoholic fermentation (*S. uvarum* and *S. cerevisiae*) were confirmed. Previous studies reported that *S. uvarum* produced higher concentrations of malic acid [42,11], succinic acid [66,67] and glycerol [11,67] that *S. cerevisiae*. In addition, this species is characterized by weaker fermentation performances than *S. cerevisiae* [11,67] and by the production of specific fermentative aromas [66,68]. Da Silva *et al.* (2015) reported that *S. uvarum* produced more acetic acid, but this finding was not supported in our study except for one strain (RC4-15) which produced 0.70 g/L of acetic acid in GR21 while the 4 other *S. uvarum* strains produced about 0.23 g/L in the same grape juice (figure S2). Using $^{1}$H-NMR-based metabolomics, fifteen non-volatile metabolites were quantified. The results obtained confirmed the differences reported for glycerol, malic acid, and succinic acid content (Figure 4). In addition, this species reveals to produce more 2,3-butanediol (in average 1.5 times more than *S. cerevisiae* strains). The amount of pyruvic acid was in average 2 times less important for *S. uvarum* than *S. cerevisiae* strains, and residual fructose was the indication of some stuck fermentations by *S. uvarum* (~10 g/L of fructose remaining).

Second, *wine* and *flor* yeasts can be distinguished by handful of compounds, mostly belonging to organic acid family (cf. figure 4 and 5). This metabolic signature could be the consequence of a differential management of central carbon metabolism as previously proposed by [44]. This might be explained by the fact that *flor* yeasts are adapted to shift their metabolism to an oxidative metabolism when sugar and nitrogen are depleted [69]. In the context of a shift diauxic during the velum formation the aptitude to consume malic acid as a secondary carbon source may constitute a selective advantage. Moreover, a correlation was found between malic acid production and succinic acid production in GR21 (no production occurred in SB19). This observation was previously reported by [70]. It is consistent with their metabolic relationships and raise the question of the metabolic pathways taken by produced and consumed malic acid. This correlation could explain a significative difference of succinic acid content between *wine* and *flor* yeasts as they already differ for malic acid consumption.

Finally, the two control strains displayed extreme differences for malic acid content, demonstrating the efficiency of the selection they came from [21,70]. They frame the natural variability of all the other strains regarding the MAC values (Figure 6A). Only some strains of *S. uvarum* can produce as much malic acid as AC1_191. On the PCA (Figure 3), the strain AC1_191 is close to *S. uvarum* group in terms of acidity management. Indeed, this strain was selected for its ability to produce important amounts of malic acid and happens to also produce important amount of succinic acid. The extreme strain AC1_191 appears to be a good strain to cope with low amounts of malic acid in musts in a context of climate change. It is able to produce malic and, secondarily, succinic acid which decrease the pH of the final wine [65]. Finally, it was able to complete the fermentation unlike some *S. uvarum* and produced low amounts of acetic acid (Figure S3) which makes it suitable for winemaking. this strain could be used for vinification itineraries aiming to conserve the freshness of white wines. The FMGS_889 is also a strain of enological interest for lowering the acidity of rich malic wines and shortening the malo-lactic fermentation of red wines [21]. The organoleptic

consequences of the use of *S. cerevisiae* strains having an opposite organic acid metabolism has been demonstrated in a previous study [57] and was not tested in the present work due to the small volume of wine fermented.

## Conclusion

A reliable and easy to apply $^{1}$H-NMR analytical method was developed to quantify the major end point metabolites of the alcoholic fermentation of grape juices. This method required a small sample size (60 µL) a basic sample dilution and a short sample analysis (5 minutes). This efficient protocol was used for evaluating the metabolomic variability of 44 yeast strains becoming to the two major species of the *Saccharomyces* genus involved in wine fermentation. Metabolomic differences between strains belonging to different genetic groups were observed suggesting that the central metabolism of distinct populations is differently regulated. Interestingly, organic acids and glycerol metabolisms constitute strongly discriminating factors. The phenotypic diversity observed impacted the final pH value of wine and was strongly correlated to the ability of strains to consume or produce malic acid.

## Fundings

- « This study received financial support from the French government in the framework of the IdEX Bordeaux University "Investments for the Future" program / GPR Bordeaux Plant Sciences ». PM received a financial support from Aquitaine lander: program MEGAVIP, 2020-2023

## Acknowledgements

This work was supported by the Bordeaux Metabolome Facility and MetaboHUB (ANR-11-INBS-0010 project).

Authors warmly thanks Prof Hervé Alexandre, Prof Joseph Schacherer and Dr Jean-Luc Legras (CIRM) for providing some of the *flor* yeast strains used in this work.

## Author contribution

Conceptualization CV and PM; Formal analysis CV and PM; Funding acquisition PM; Investigation CV, MB, MM, GDC, ILM; Resources YN, PM; Software CV, PM; Supervision PM, TR; Writing – original draft CV, PM; Writing – review & editing CV, PM, GDC, TR

## References

1. Tamang JP, Watanabe K, Holzapfel WH. Review: Diversity of microorganisms in global fermented foods and beverages. Front Microbiol. 2016;7. doi:10.3389/fmicb.2016.00377

2. Bourdichon F, Casaregola S, Farrokh C, Frisvad JC, Gerds ML, Hammes WP, *et al.* Food

522 fermentations: Microorganisms with technological beneficial use. Int J Food Microbiol. 2012;154: 87–97.
523 doi:10.1016/j.ijfoodmicro.2011.12.030
524 3. Ebeler SE, Thorngate JH. Wine chemistry and flavor: Looking into the crystal glass. J Agric Food
525 Chem. 2009;57: 8098–8108. doi:10.1021/jf9000555
526 4. Alañón ME, Pérez-Coello MS, Marina ML. Wine science in the metabolomics era. TrAC Trends Anal
527 Chem. 2015;74: 1–20. doi:https://doi.org/10.1016/j.trac.2015.05.006
528 5. Pinu FR. Grape and Wine Metabolomics to Develop New Insights Using Untargeted and Targeted
529 Approaches. Fermentation. 2018. doi:10.3390/fermentation4040092
530 6. Lorrain B, Ky I, Pechamat L, Teissedre P-L. Evolution of Analysis of Polyhenols from Grapes, Wines,
531 and Extracts. Molecules. 2013. pp. 1076–1100. doi:10.3390/molecules18011076
532 7. Fleet GH. Wine yeasts for the future. FEMS Yeast Res. 2008;8: 979–995. doi:10.1111/j.1567-
533 1364.2008.00427.x
534 8. Ribéreau-Gayon, P., Dubourdien, D., Donéche B& AL. Handbook of Enology Volume 1 The
535 Microbiology of Wine and Vinifications 2 nd Edition Handbook of Enology Volume 1 The Microbiology of
536 Wine and Vinifications 2 nd Edition. 2006.
537 9. Swiegers JH, Bartowsky EJ, Henschke PA, Pretorius IS. Yeast and bacterial modulation of wine aroma
538 and flavour. Aust J Grape Wine Res. 2005;11: 139–173. doi:10.1111/j.1755-0238.2005.tb00285.x
539 10. Hranilovic A, Gambetta JM, Schmidtke L, Boss PK, Grbin PR, Masneuf-Pomarede I, *et al*. Oenological
540 traits of *Lachancea thermotolerans* show signs of domestication and allopatric differentiation. Sci Rep. 2018;8:
541 1–13. doi:10.1038/s41598-018-33105-7
542 11. Masneuf-Pomarède I, Bely M, Marullo P, Lonvaud-Funel A, Dubourdieu D. Reassessment of
543 phenotypic traits for *Saccharomyces bayanus var. uvarum* wine yeast strains. Int J Food Microbiol. 2010;139:
544 79–86. doi:10.1016/j.ijfoodmicro.2010.01.038
545 12. Da Silva T, Albertin W, Dillmann C, Bely M, La Guerche S, Giraud C, *et al.* Hybridization within
546 *Saccharomyces* genus results in homoeostasis and phenotypic novelty in winemaking conditions. PLoS One.
547 2015;10: 1–24. doi:10.1371/journal.pone.0123834
548 13. Viana F, Gil J V, Genovés S, Vallés S, Manzanares P. Rational selection of non-*Saccharomyces* wine
549 yeasts for mixed starters based on ester formation and enological traits. Food Microbiol. 2008;25: 778–785.
550 doi:https://doi.org/10.1016/j.fm.2008.04.015
551 14. Soles RM, Ough CS, Kunkee RE. Ester Concentration Differences in Wine Fermented by Various
552 Species and Strains of Yeasts. Am J Enol Vitic. 1982;33: 94–98.
553 15. Chatonnet P, Dubourdie D, Boidron J, Pons M. The origin of ethylphenols in wines. J Sci Food Agric.
554 1992;60: 165–178. doi:https://doi.org/10.1002/jsfa.2740600205
555 16. Petitgonnet C, Klein GL, Roullier-Gall C, Schmitt-Kopplin P, Quintanilla-Casas B, Vichi S*, et al.*
556 Influence of cell-cell contact between *L. thermotolerans* and *S. cerevisiae* on yeast interactions and the exo-
557 metabolome. Food Microbiol. 2019;83: 122–133. doi:10.1016/j.fm.2019.05.005
558 17. Howell KS, Cozzolino D, Bartowsky EJ, Fleet GH, Henschke P a. Metabolic profiling as a tool for
559 revealing *Saccharomyces* interactions during wine fermentation. FEMS Yeast Res. 2006;6: 91–101.
560 doi:10.1111/j.1567-1364.2005.00010.x
561 18. Renault P, Coulon J, Moine V, Thibon C, Bely M. Enhanced 3-Sulfanylhexan-1-ol Production in

562 Sequential Mixed Fermentation with *Torulaspora delbrueckii/Saccharomyces cerevisiae* Reveals a Situation of
563 Synergistic Interaction between Two Industrial Strains. Frontiers in Microbiology. 2016. p. 293.
564 19. Tempère S, Marchal A, Barbe J-C, Bely M, Masneuf-Pomarede I, Marullo P, *et al*. The complexity of
565 wine: clarifying the role of microorganisms. Appl Microbiol Biotechnol. 2018;102: 3995–4007.
566 doi:10.1007/s00253-018-8914-8
567 20. Tilloy V, Ortiz-Julien A, Dequin S. Reduction of ethanol yield and improvement of glycerol formation
568 by adaptive evolution of the wine yeast *Saccharomyces cerevisiae* under hyperosmotic conditions. Appl Environ
569 Microbiol. 2014;80: 2623–32. doi:10.1128/AEM.03710-13
570 21. Vion C, Peltier E, Bernard M, Muro M, Marullo P. Marker assisted selection of malic-consuming
571 *Saccharomyces cerevisiae* strains for winemaking. Efficiency and limits of a qtl's driven breeding program. J
572 Fungi. 2021;7. doi:10.3390/jof7040304
573 22. Dufour M, Zimmer A, Thibon C, Marullo P. Enhancement of volatile thiol release of *Saccharomyces*
574 *cerevisiae* strains using molecular breeding. Appl Microbiol Biotechnol. 2013;97: 5893–5905.
575 doi:10.1007/s00253-013-4739-7
576 23. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, *et al.* Genome evolution across
577 1,011 *Saccharomyces cerevisiae* isolates Species-wide genetic and phenotypic diversity. Nature. 2018;556: 339–
578 347.
579 24. Steyer D, Ambroset C, Brion C, Claudel P, Delobel P, Sanchez I, *et al*. QTL mapping of the production
580 of wine aroma compounds by yeast. BMC Genomics. 2012;13: 573. doi:10.1186/1471-2164-13-573
581 25. Eder M, Sanchez I, Brice C, Camarasa C, Legras JL, Dequin S. QTL mapping of volatile compound
582 production in *Saccharomyces cerevisiae* during alcoholic fermentation. BMC Genomics. 2018;19: 1–19.
583 doi:10.1186/s12864-018-4562-8
584 26. Peltier E, Sharma V, Martí Raga M, Roncoroni M, Bernard M, Jiranek V, *et al.* Dissection of the
585 molecular bases of genotype x environment interactions: A study of phenotypic plasticity of *Saccharomyces*
586 *cerevisiae* in grape juices. BMC Genomics. 2018;19: 1–20. doi:10.1186/s12864-018-5145-4
587 27. Cubillos FA, Brice C, Molinet J, Tisné S, Abarca V, Tapia SM, *et al*. Identification of Nitrogen
588 Consumption Genetic Variants in Yeast Through QTL Mapping and Bulk Segregant RNA-Seq Analyses. G3
589 Genes, Genomes, Genet. 2017;7: 1693–1705. doi:10.1534/g3.117.042127
590 28. Bony M, Bidart F, Camarasa C, Ansanay V, Dulau L, Barre P, *et al*. Metabolic analysis of *S. cerevisiae*
591 strains engineered for malolactic fermentation. FEBS Lett. 1997;410: 452–456. doi:10.1016/S0014-
592 5793(97)00637-6
593 29. Bordet F, Roullier-Gall C, Ballester J, Vichi S, Quintanilla-Casas B, Gougeon RD, *et al*. Different
594 wines from different yeasts? "*Saccharomyces cerevisiae* intraspecies differentiation by metabolomic signature
595 and sensory patterns in wine." Microorganisms. 2021;9: 1–17. doi:10.3390/microorganisms9112327
596 30. Reed LK, Baer CF, Edison AS. Considerations when choosing a genetic model organism for
597 metabolomics studies. Curr Opin Chem Biol. 2017;36: 7–14. doi:10.1016/j.cbpa.2016.12.005
598 31. Cusano E, Cagliani LR, Consonni R, Simonato B, Zapparoli G. NMR-based metabolic profiling of
599 different yeast fermented apple juices. Lwt. 2020;118. doi:10.1016/j.lwt.2019.108771
600 32. Avenoza A, Busto JH, Canal N, Peregrina JM. Time course of the evolution of malic and lactic acids in
601 the alcoholic and malolactic fermentation of grape must by quantitative 1H NMR (qHNMR) spectroscopy. J

602 Agric Food Chem. 2006;54: 4715–4720. doi:10.1021/jf060778p
603 33. Emwas AH, Roy R, McKay RT, Tenori L, Saccenti E, Nagana Gowda GA, *et al*. NMR Spectroscopy
604 for Metabolomics Research. Metabolites. 2019;9. doi:10.3390/metabo9070123
605 34. Puig-Castellví F, Alfonso I, Piña B, Tauler R. 1H NMR metabolomic study of auxotrophic starvation in
606 yeast using Multivariate Curve Resolution-Alternating Least Squares for Pathway Analysis. Sci Rep. 2016;6:
607 30982. doi:10.1038/srep30982
608 35. Vignoli A, Ghini V, Meoni G, Licari C, Takis PG, Tenori L, *et al.* High-Throughput Metabolomics by
609 1D NMR. Angew Chemie - Int Ed. 2019;58: 968–994. doi:10.1002/anie.201804736
610 36. Dequin S, Escudier J-L, Bely M, Noble J, Albertin W, Masneuf-Pomarède I, *et al.* How to adapt
611 winemaking practices to modified grape composition under climate change conditions. OENO One. 2017;51:
612 205. doi:10.20870/oeno-one.2016.0.0.1584
613 37. Volschenk H, van Vuuren HJJ, Viljoen-Bloom M. Malic Acid in Wine: Origin, Function and
614 Metabolism during Vinification. South African J Enol Vitic. 2006;27. doi:10.21548/27-2-1613
615 38. Van Leeuwen C, Destrac-Irvine A. Modified grape composition under climate change conditions
616 requires adaptations in the vineyard. Oeno One. 2017;51: 147–154. doi:10.20870/oeno-one.2016.0.0.1647
617 39. Bigard A, Berhe DT, Maoddi E, Sire Y, Boursiquot JM, Ojeda H, *et al*. Vitis vinifera L. fruit diversity
618 to breed varieties anticipating climate changes. Front Plant Sci. 2018;9: 0–16. doi:10.3389/fpls.2018.00455
619 40. Yéramian N, Chaya C, Suárez Lepe JA. L-(-)-malic acid production by *Saccharomyces* spp. during the
620 alcoholic fermentation of wine. J Agric Food Chem. 2007;55: 912–919. doi:10.1021/jf061990w
621 41. Gao C, Fleet GH. Degradation of malic and tartaric acids by high density cell suspensions of wine
622 yeasts. Food Microbiol. 1995.
623 42. Coloretti F, Zambonelli C, Castellari L, Tini V, Rainieri S. The effect of DL-malic acid on the
624 metabolism of L-malic acid during wine alcoholic fermentation. Food Technol Biotechnol. 2002;40: 317–320.
625 43. Peltier E, Bernard M, Trujillo M, Barbe J-C, Prodhomme D, Gibon Y, *et al*. Wine yeast phenomics: A
626 standardized fermentation method for assessing quantitative traits of *Saccharomyces cerevisiae* strains in
627 enological conditions. PLoS One. 2018; 1–23. doi:10.1371/journal.pone.0190094
628 44. Peltier E, Vion C, Abou Saada O, Friedrich A, Schacherer J, Marullo P. Flor Yeasts Rewire the Central
629 Carbon Metabolism During Wine Alcoholic Fermentation. Front Fungal Biol. 2021;2.
630 doi:10.3389/ffunb.2021.733513
631 45. Le Mao I, Martin-Pernier J, Bautista C, Lacampagne S, Richard T, Da Costa G. 1H-Nmr Metabolomics
632 as a Tool for Winemaking Monitoring. Molecules. 2021;26: 1–13. doi:10.3390/molecules26226771
633 46. Gougeon L, da Costa G, Guyon F, Richard T. 1H NMR metabolomics applied to Bordeaux red wines.
634 Food Chem. 2019;301: 125257. doi:10.1016/j.foodchem.2019.125257
635 47. Albertin W, Da Silva T, Rigoulet M, Masneuf- Pomarede I, De Vienne D, Sicard D, *et al.* The
636 Mitochondrial Genome Impacts Respiration but Not Fermentation in Interspecific *Saccharomyces* Hybrids.
637 PLoS One. 2013.
638 48. Peltier E, Bernard M, Trujillo M, Prodhomme DD, Barbe J-C, Gibon Y, *et al*. Wine yeast phenomics: a
639 standardized fermentation method for assessing quantitative traits of *Saccharomyces cerevisiae* strains in
640 enological conditions. Schacherer J, editor. PLoS One. 2018;13: 191353. doi:10.1101/191353
641 49. Albertin W, Chernova M, Durrens P, Guichoux E, Sherman DJ, Masneuf-Pomarede I, *et al*. Many

642 interspecific chromosomal introgressions are highly prevalent in Holarctic *Saccharomyces uvarum* strains found
643 in human-related fermentations. Yeast. 2017;35: 141–156. doi:10.1002/yea.3248
644 50. Börlin M, Venet P, Claisse O, Salin F, Legras JL, Masneuf-Pomarede I. Cellar-associated
645 *Saccharomyces cerevisiae* population structure revealed high-level diversity and perennial persistence at
646 Sauternes wine estates. Appl Environ Microbiol. 2016;82: 2909–2918. doi:10.1128/AEM.03627-15
647 51. Airoldi C, Tripodi F, Guzzi C, Nicastro R, Coccetti P. NMR analysis of budding yeast metabolomics: A
648 rapid method for sample preparation. Mol Biosyst. 2015;11: 379–383. doi:10.1039/c4mb00452c
649 52. Cobas C, Seoane F, Domínguez S, Sykora S. A new approach to improving automated analysis of
650 proton NMR spectra through Global Spectral Deconvolution (GSD). Spectrosc Eur. 2011;23: 26–30.
651 53. R Development Core Team. R: A language and environment for statistical computing. 2016;2: 1–12.
652 54. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A*, et al.* Genome evolution across
653 1,011 *Saccharomyces cerevisiae* isolates. Nature. 2018;556: 339–344. doi:10.1038/s41586-018-0030-5
654 55. Coi AL, Bigey F, Mallet S, Marsit S, Zara G, Gladieux P*, et al.* Genomic signatures of adaptation to
655 wine biological aging conditions in biofilm-forming flor yeasts. Mol Ecol. 2017. doi:10.1111/mec.14053
656 56. Chernova M, Albertin W, Durrens P, Guichoux E, Sherman DJ, Masneuf-Pomarede I, *et al*. Many
657 interspecific chromosomal introgressions are highly prevalent in Holarctic *Saccharomyces uvarum* strains found
658 in human-related fermentations. Yeast. 2018;35: 141–156. doi:10.1002/yea.3248
659 57. Vion C, Muro M, Bernard, Margaux, Richard B, Fautre V, Yeramian N, *et al.* New malic acid producer
660 strains of *Saccharomyces cerevisiae* for preserving wine acidity during alcoholic fermentation. Food Microbiol.
661 2022;112. doi:10.1016/j.fm.2022.104209
662 58. Crépin L, Nidelet T, Sanchez I, Dequin S, Camarasa C. Sequential use of nitrogen compounds by
663 *Saccharomyces cerevisiae* during wine fermentation: A model based on kinetic and regulation characteristics of
664 nitrogen permeases. Appl Environ Microbiol. 2012;78: 8102–8111. doi:10.1128/AEM.02294-12
665 59. Park YC, Shaffer CEH, Bennett GN. Microbial formation of esters. Appl Microbiol Biotechnol.
666 2009;85: 13–25. doi:10.1007/s00253-009-2170-x
667 60. Calull M, Marcé RM, Borrull F. Determination of carboxylic acids, sugars, glycerol and ethanol in wine
668 and grape must by ion-exchange high-performance liquid chromatography with refractive index detection. J
669 Chromatogr A. 1992;590: 215–222. doi:10.1016/0021-9673(92)85384-6
670 61. Lehtonen P. Determination of Amines and Amino Acids in Wine — A Review. Am J Enol Vitic.
671 1996;47: 127 LP – 133. doi:10.5344/ajev.1996.47.2.127
672 62. Godelmann R, Kost C, Patz CD, Ristow R, Wachter H. Quantitation of compounds in wine using 1H
673 NMR spectroscopy: Description of the method and collaborative study. J AOAC Int. 2016;99: 1295–1304.
674 doi:10.5740/jaoacint.15-0318
675 63. Nascimento ESP, Cardoso DR, Franco DW. Quantitative ester analysis in cachaça and distilled spirits
676 by Gas Chromatography-Mass Spectrometry (GC-MS). J Agric Food Chem. 2008;56: 5488–5493.
677 doi:10.1021/jf800551d
678 64. Alcarde AR, Souza LM, Bortoletto AM. Formation of volatile and maturation-related congeners during
679 the aging of sugarcane spirit in oak barrels. J Inst Brew. 2014;120: 529–536. doi:10.1002/jib.165
680 65. Liu SQ, Pilone GJ. An overview of formation and roles of acetaldehyde in winemaking with emphasis
681 on microbiological implications. Int J Food Sci Technol. 2000;35: 49–61. doi:10.1046/j.1365-

2621.2000.00341.x

66. Pérez D, Jaehde I, Guillamón JM, Heras JM, Querol A. Screening of *Saccharomyces* strains for the capacity to produce desirable fermentative compounds under the influence of different nitrogen sources in synthetic wine fermentations. Food Microbiol. 2021;97. doi:10.1016/j.fm.2021.103763

67. Minebois R, Pérez-Torrado R, Querol A. A time course metabolism comparison among *Saccharomyces cerevisiae, S. uvarum and S. kudriavzevii* species in wine fermentation. Food Microbiol. 2020;90: 103484. doi:10.1016/j.fm.2020.103484

68. Pérez D, Denat M, Heras JM, Guillamón JM, Ferreira V, Querol A. Effect of non-wine *Saccharomyces* yeasts and bottle aging on the release and generation of aromas in semi-synthetic Tempranillo wines. Int J Food Microbiol. 2022;365. doi:10.1016/j.ijfoodmicro.2022.109554

69. Legras JL, Moreno-Garcia J, Zara S, Zara G, Garcia-Martinez T, Mauricio JC, *et al.* Flor yeast: New perspectives beyond wine aging. Front Microbiol. 2016;7: 1–11. doi:10.3389/fmicb.2016.00503

70. Vion C, Muro M, Bernard, Margaux, Richard B, Fautre V, Yeramian N*, et al.* New malic acid producer strains of *Saccharomyces cerevisiae* for preserving wine acidity during alcoholic fermentation. Food Microbiol. 2022.

71. David-Vaizant V, Alexandre H. Flor yeast diversity and dynamics in biologically aged wines. Front Microbiol. 2018;9: 1–16. doi:10.3389/fmicb.2018.02235

72. Huang C, Roncoroni M, Gardner RC. MET2 affects production of hydrogen sulfide during wine fermentation. Appl Microbiol Biotechnol. 2014;98: 7125–7135. doi:10.1007/s00253-014-5789-1

73. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: a R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014;2: e281. doi:10.7717/peerj.281

74. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 2004;20: 289–290. doi:10.1093/bioinformatics/btg412

# Figures

All figures should be printed in color

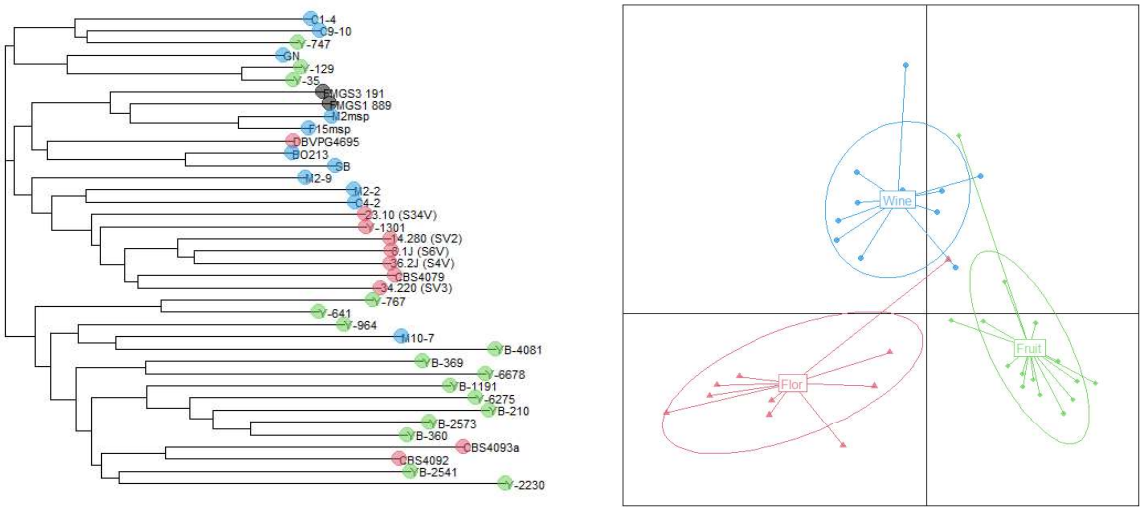



**Figure 1:** (A) Genetic relationships between the 37 S. cerevisiae strains of the study (Bruvo's distance) computed from the genotyping of 14 variable microsatellites, the colors red, green, and blue indicated the Flor, Fruit, and Wine origin, respectively. the two control strains are indicated by a black dot (b) Discriminant Analysis of Principal Components computed with microsatellite data. The three main populations are figured out by the same colors.

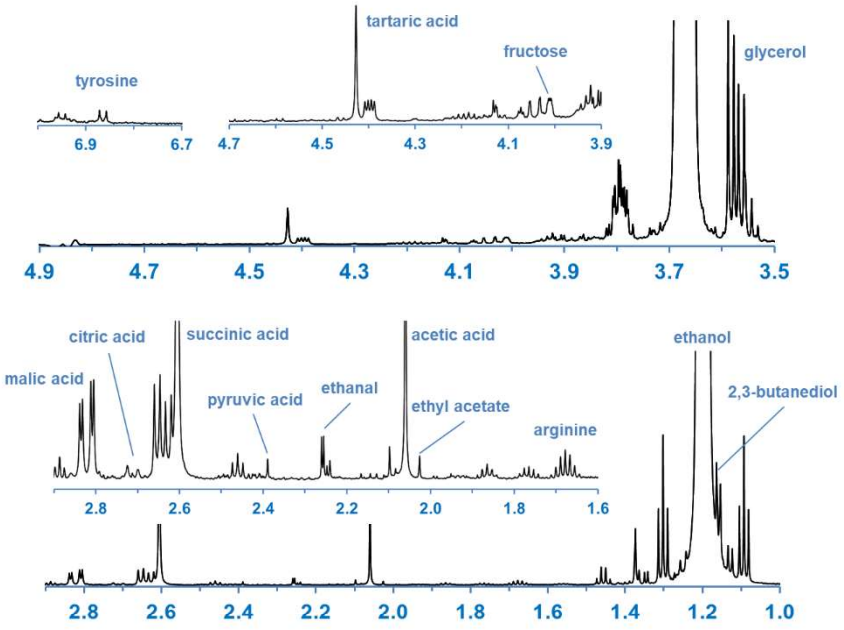
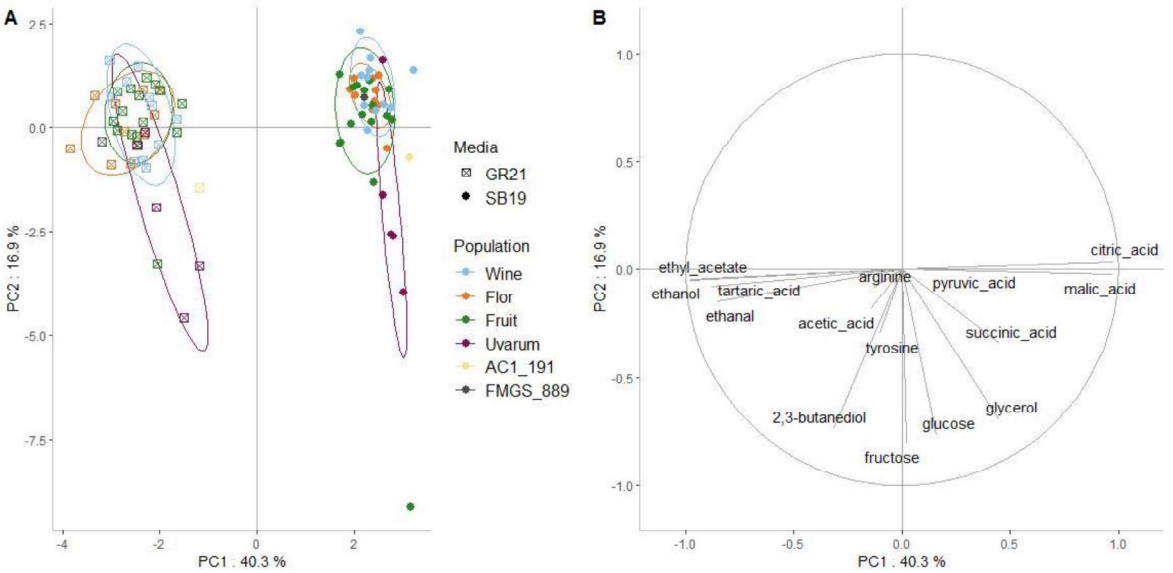



**Figure 2:** Annotated typical $^{1}$H-NMR spectrum of wine metabolites after water suppression (noesygppr1d). Identified constituents are listed in Table 3.

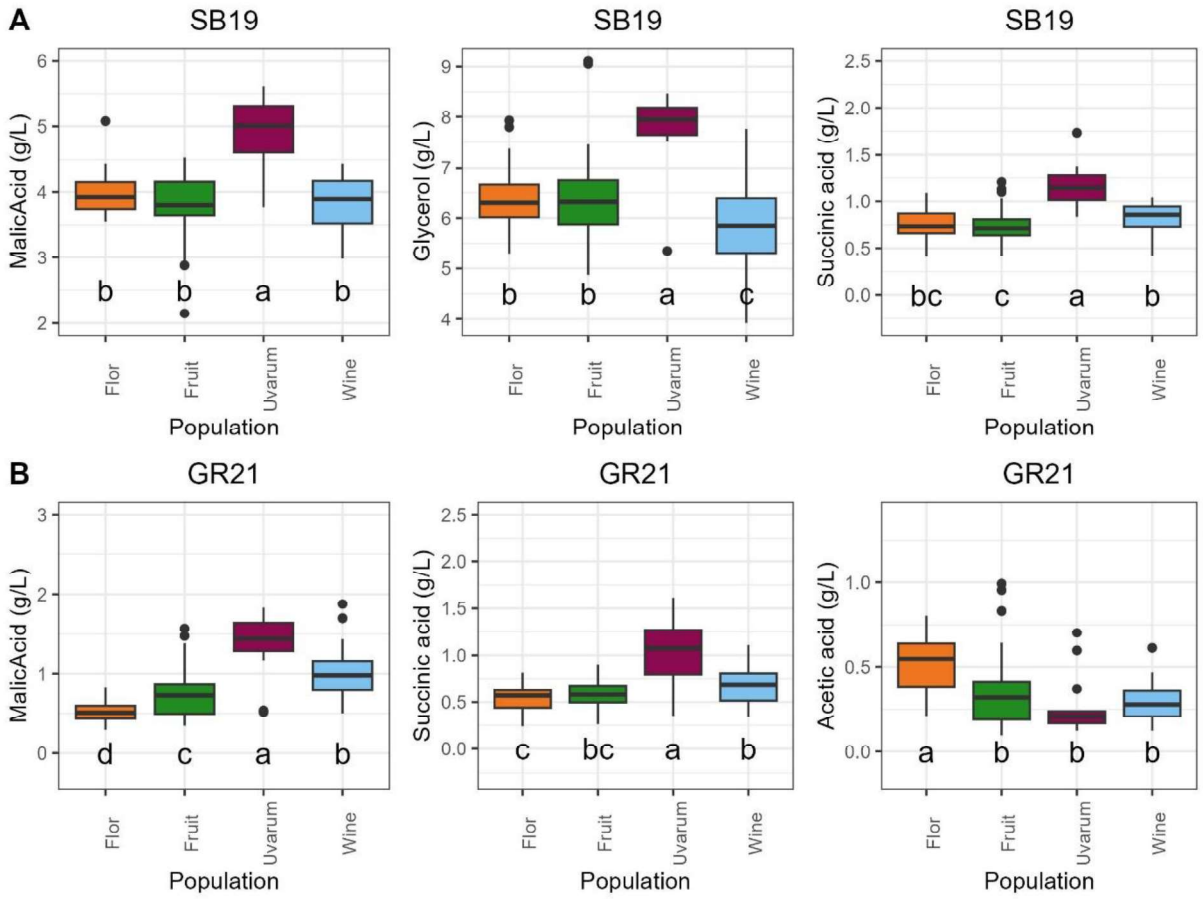



717

718 **Figure 3:** Principal Component Analysis (PCA) (A) The two first axis of the PCA performed for the 44 strains in the two
719 grape juices. Axes 1 and 2 explain 40.3 and 16.9 % of total variation respectively. Each point represents the average of three
720 biological replicates of a single strain. Each strain is colored according to its population. (B) Correlation circle indicating the
721 correlation of the variables for axes 1 and 2.

722

723




724

725 **Figure 4:** Main compounds involved in populations separation. (**A**) Absolute concentrations (g/L) of malic acid, glycerol,
726 and succinic acid, in SB19 for the 4 populations. (**B**) Absolute concentrations (g/L) of malic acid, succinic acid, and acetic

acid in GR21 for the four yeast populations. A Kruskal test has been applied to the compounds in both media. Different letters indicate a significant difference between the populations ($\alpha < 0.05$).

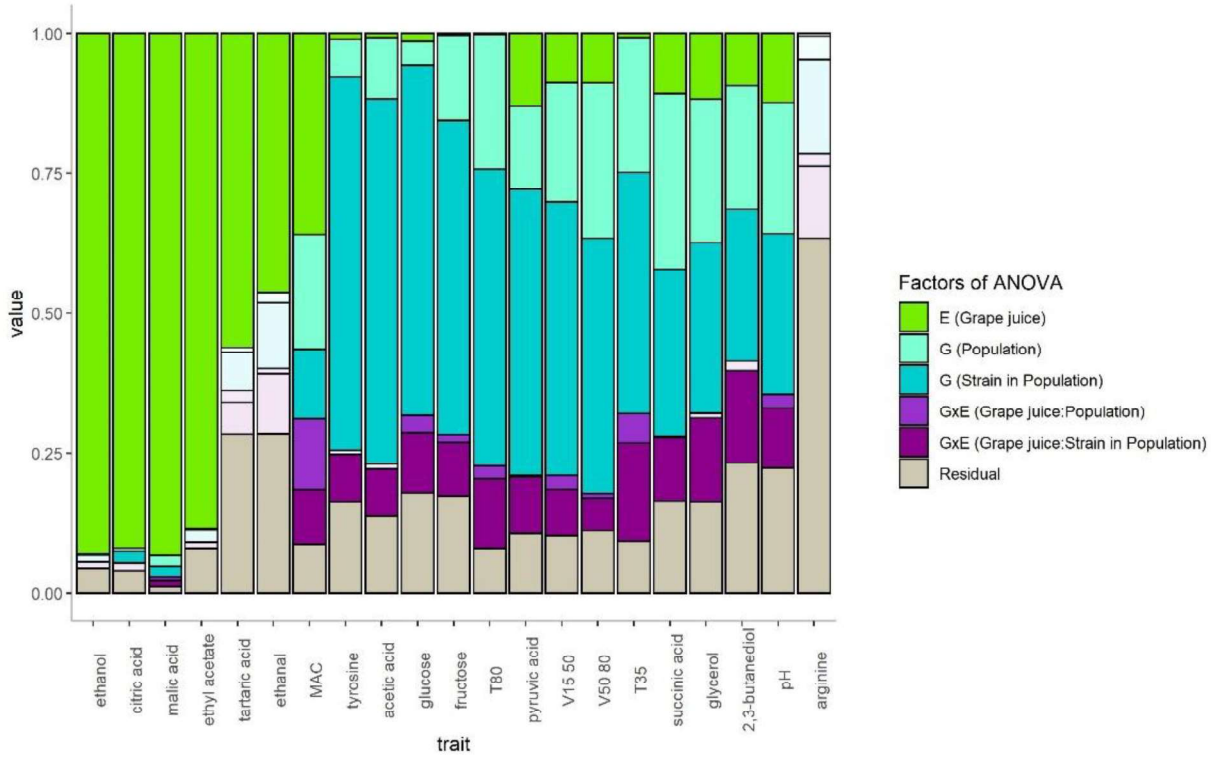



**Figure 5:** Bar graphs indicating the part of variance explained by the different factors of the ANOVA. The letters E and G represent the environmental (grape juice) and the genetic factors, respectively. The nested ANOVA applied allows to evaluate the effect of the population and the effect of strain within each population (strain). GxE represents the interaction between genetic and environment and was decomposed in two nested factors grape juice: population and grape juice: strain in population. Gray tons indicate non significative effect of the factors.

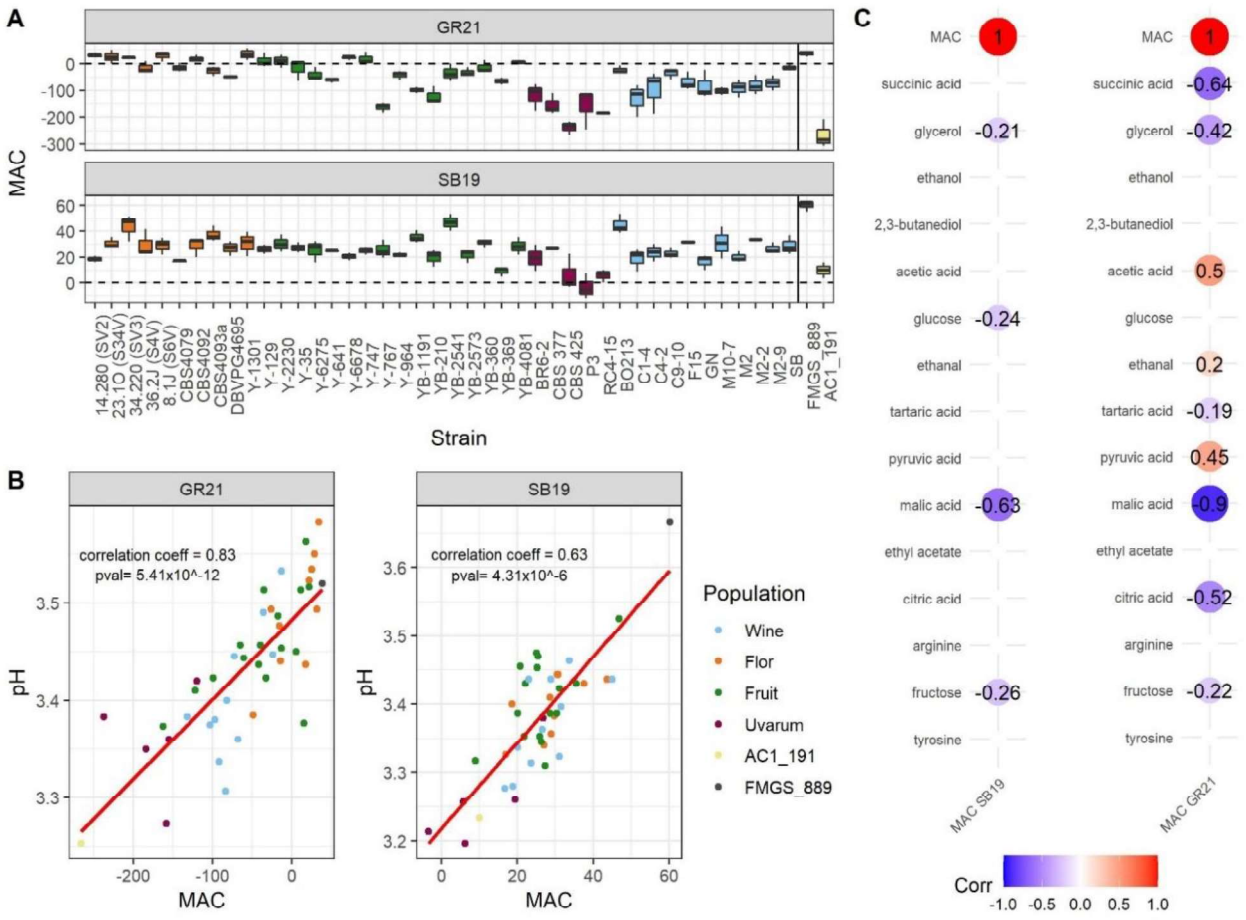



**Figure 6:** (**A**) MAC values measured for all strains in the two grape juices (3 replicates per strain). Strains were colored according to their population. The dashed line represents the 0% which corresponds to no malic acid consumed or produced. Positive values correspond to a consumption of malic acid and negative values to a production. (**B**) Correlation of MAC and pH for the two grape juices. The red line represents the linear regression line, correlation coefficients and p-values are written on the graphs. The dots represent all the tested strains (3 replicates per strains). They are colored according to their genetic group. (**C**) Correlation matrix between MAC and all other metabolites in the two grape juices. The value displayed corresponds to the correlation coefficient. Red values correspond to a negative correlation while blue values correspond to a positive correlation, only significative correlations ($\alpha < 0.05$) were indicated.

# Supplementary data

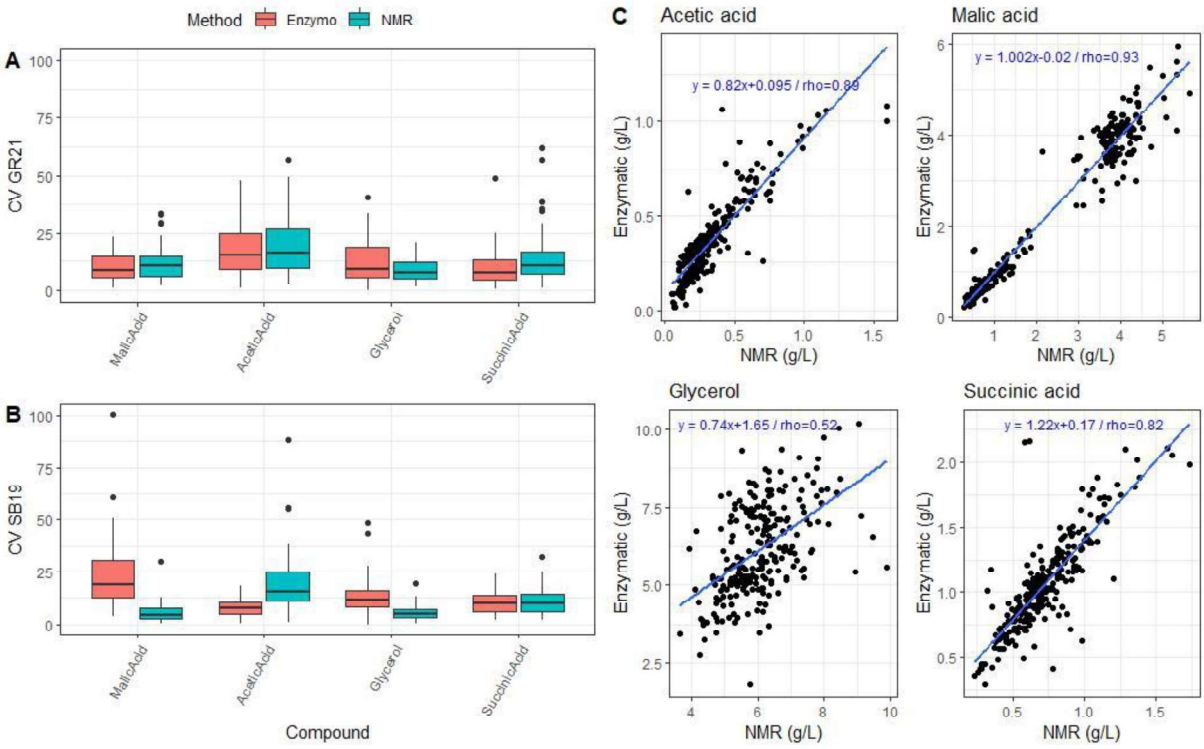



**Figure S1:** Correlation between enzymatic assay and NMR analysis for acetic acid, glycerol, malic acid, and succinic acid. The Spearman correlation coefficients are 0.89, 0.82, 0.93 and 0.52 for acetic acid, succinic acid, malic acid, and glycerol, respectively. The concentrations on both axes are expressed in g/L. There is an overestimation of the concentration of malic acid, succinic acid, and glycerol with the NMR method.

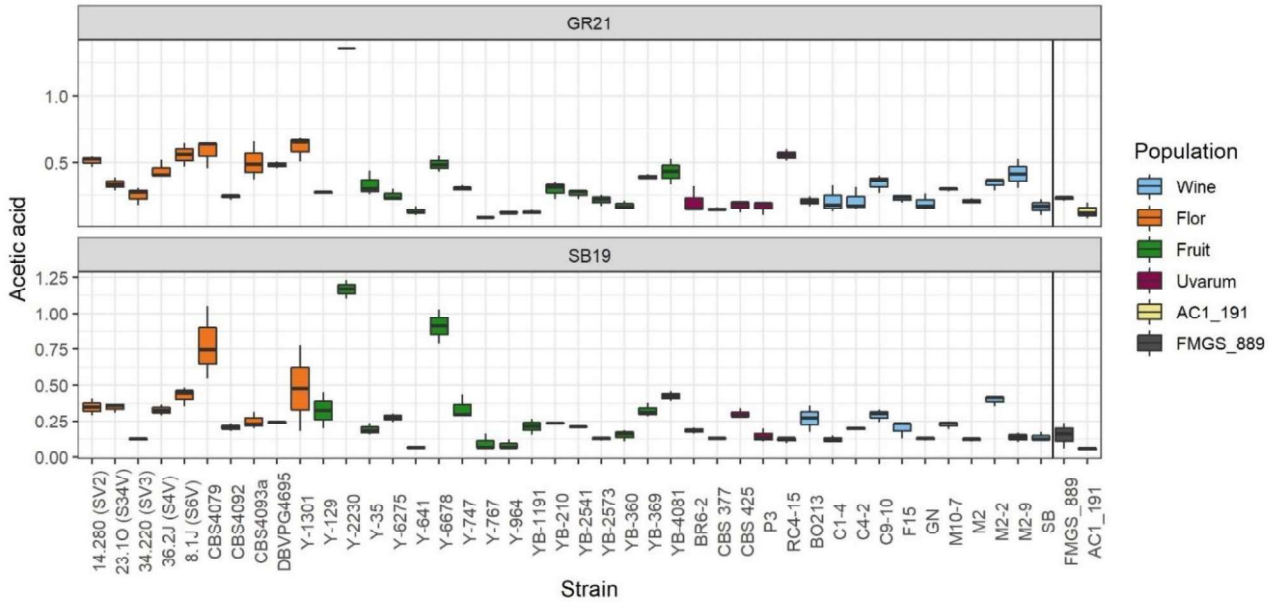



**Figure S2:** Absolute concentration in acetic acid (g/L) at the end of the fermentation for all strains in both grape juices. The boxplots are colored according to their population.