Multi locus sequence typing of oenological *Saccharomyces cerevisiae* strains.

Rosario Muñoz, Alicia Gomez, Virginia Robles, Patricia Rodriguez, Eduardo Cebollero, Laura Tabera, Alfonso V. Carrascosa, Ramón Gonzalez*

Departamento de Microbiología, Instituto de Fermentaciones Industriales (CSIC), Madrid, Spain.

*Corresponding author. Present address. Instituto de Ciencias de la Vid y el Vino (CSIC/Universidad de La Rioja/Gobierno de La Rioja)

ABSTRACT.

Wine yeast strains, isolated from diverse Spain wine producing areas, and molecularly characterized as *Saccharomyces cerevisiae*, as well as some commercial wine yeast strains, were typed by two alternative molecular methods, the well established mitochondrial RFLP analysis, and by a multilocus sequence typing (MLST) designed scheme. The discrimination potential of mitochondrial RFLP analysis was superior to the MLST scheme used in this work. Ten polymorphic sites were found in the five nuclear loci analyzed showing 13 different genotypes, with 11 of them represented by only one strain. However, MLST analysis allowed easy construction of reliable phylogenetic trees. Although by MLST analysis, wine isolates of *S. cerevisiae* appeared as a rather homogeneous group, split decomposition analysis indicated that recombination plays a role in creating some genetic heterogeneity in wine *S. cerevisiae* strains. These results are in contrast to the genetic diversity revealed by MLST in other yeast species, like *Candida albicans*.

Keywords: wine yeast, MLST, genetic typing

1. INTRODUCTION

The main species responsible for alcoholic fermentation in winemaking is *Saccharomyces cerevisie* (Querol and Fleet, 2006). Apart from alcoholic fermentation, yeast cells contribute to wine attributes trough the release of low molecular weight molecules like glycerol, acetate, succinate, pyruvate, and several esters, all of them contributing to the sensorial properties of the wine (Fleet, 1993). In addition, yeast cells release cell constituents like proteins or polysaccharides, also involved in wine quality (Escot et al., 2001). Traditional wine making is based on the spontaneous fermentation of grape must by indigenous yeasts, a variable mixture of strains belonging to several genera that, together with viticultural and technological aspects might contribute to the "terroir" character of some wines. However the advantages of spontaneous fermentations are often outweighed by the risk of sluggish, stuck, or defective fermentations from a sensorial perspective. To avoid these troubles winemakers use starter cultures of selected yeast strains, usually commercialized as active dry yeast. Nowadays the yeast strain factor is recognized as having a major influence in the quality of both still and sparkling wines (Kunkee and Amerine, 1970; Querol and Ramon, 1996; Martinez-Rodriguez et al., 2001). The particular strain used should be adapted to the particular winemaking style. Selection criteria for wine yeast strains have been discussed in several reviews (Fleet, 1993); and have evolved from assuring complete fermentation with suitable kinetics; to quality related properties, like production or release of primary and secondary aroma compounds, other sensory properties, tolerance to difficult fermentation conditions, killer phenotype, chemical stability of the wine produced, or technological properties.

In this context, the DNA molecular typing of yeast strains has several applications, including monitoring the dominance of the inoculated yeast strain, yeast population dynamics studies, studies of wine yeast strain origin and evolution, and protection of the industrial property on commercial yeast strains (Querol et al., 1992; Guillamon et al., 1998; Fernandez-Espinar *et* al., 2001; Torija et al., 2001). Available molecular typing techniques that have been applied to the genetic identification of wine yeast strains include, separation of intact chromosomes by pulsed field agarose gel electrophoresis (Vezinhet et al., 1990; Guilllamon et al., 1998); restriction analysis of the mitochondrial genome (Vezinhet et al., 1990; Querol et al., 1992); analysis of δ sequences by PCR amplification (Ness et al., 1993; Lavalle et al., 1994; Legras and Karst, 2003); microsatellite markers (Balerias Couto et al., 1996; Techera et al., 2001; Gallego et al., 1998), PCR amplification of the mitochondrial, intron rich, *COX1* gene (Lopez et al., 2003), Random Amplified Polimorfic DNA (RAPD-PCR) (Grando et al., 1994; Quesada and Cenis, 1995), or combination of several of these methodologies (Fernandez-Espinar *et* al., 2001).

Multilocus sequence typing (MLST) was recently shown to be a powerful technique for typing microorganisms. Typically, different alleles of housekeeping genes are characterized by PCR amplification and automated sequencing (Enright and Spratt, 1999). It is highly discriminatory at the strain level and, being based in nucleotide sequencing, the results are easily comparable between laboratories. It has been applied in the fields of clinical epidemiology (Urwin and Maiden, 2003; Enright and

Spratt, 1999; Maiden et al., 1998; Bougnoux et al., 2002; Sullivan et al., 2005), microbial food safety (Dingle et al., 2001; Farfan et al., 2002; Helgason et al., 2004), and food biotechnology (De las Rivas et al., 2004; De las Rivas et al., 2006).

More recently, multi locus sequence typing has been applied to *S. cerevisiae* (Fay and Benavides, 2005), and there are just a few reports on its use for typing wine yeast strains. In this work we used MLST analysis for molecular typing of several *S. cerevisiae* wine yeast strains, including isolates from sherry and sparkling wines, and commercial strains. The MLST discrimination power is discussed in comparison to mitochondrial RFLP analysis.

2. MATERIALS AND METHODS

2.1. Yeast strains

All yeast strains used in this work are listed in Table 1. All them were previously classified as *S. cerevisiae* by phenotypic analysis and by RFLP analysis of 5.8S-ITS region (Fernandez-Espinar *et* al., 2000) with the restriction enzymes *Cfo*I*,* *Hae*III and *Scr*FI. The pattern of most strains was the most common *S. cerevisiae* pattern, as established by (Fernandez-Espinar *et* al., 2000), labeled as pattern A in Table 1, but some of the strains isolated from “wine flor” showed the specific pattern for *S. cerevisiae* flor yeast strains, labeled as pattern B in Table 1.

2.2. Mitochondrial RFLP analysis

Mitochondrial DNA restriction patterns were obtained by the method of Querol et al., (1992) by using the restriction endonuclease *Hin*fI. Briefly, DNA was purified as described by Querol et al., (1992) and digested with the restriction enzyme *Hin*fI (Roche Diagnostics, Barcelona, Spain) following the instructions of the supplier. Restriction fragments were separated by electrophoresis in 1 % (w/v) agarose gels in 1XTAE buffer, stained with ethidium bromide (0.5 µg/ml). DNA from phage λ digested with *Eco*RI and *Hin*dIII endonucleases was used as molecular weight marker. The images were visualized on a 312 nm UV Transilluminator and recorded with a Digi Doc Documentation System (Bio-Rad Laboratories, Madrid, Spain).

2.3. MLST analysis

2.3.1. PCR amplification and DNA sequencing

The genes finally chosen for the MLST analysis encoded the following proteins: a putative ATP-dependent permease of the ABC transporter family (ADP1p), acetyl-CoA carboxylase (ACC1p), a subunit of the 26S proteasome (RPN2p), glutamine tRNA synthetase (GLN4p), and alanyl-tRNA synthetase (ALA1p). The DNA sequences of these candidate loci are available from the GeneBank. Gene fragments were amplified by PCR from chromosomal DNA of the wine *S. cerevisiae* strains. Sequence of the primers used for the amplification of the gene fragments are shown in Table 2. The conditions of PCR, purification, and DNA sequencing were previously described (De las Rivas et al., 2004).

2.3.2. Data analysis

For each locus, the sequences obtained for all isolates were compared, and the different sequences were assigned arbitrary allele numbers. For each isolate, the combination of genotypes obtained at each locus defined its genotype profile. Each isolate was therefore designated by five numbers, constituting a diploid sequence type (DST) (Table 1).

Sequence alignments and comparison were done with the program BioEdit (http://jwbrown.mbio.ncsu.edu/BioEdit/bioedit.html) and converted into MEGA and NEXUS files with START. Phylogenetic tree was showed as a cladogram. A cladogram is a branching diagram (tree) assumed to be an estimate of a phylogeny where the branches are of equal length, thus cladograms show common ancestry, but do not indicate the amount of evolutionary "time" separating taxa.

The method of split decomposition was used to assess the degree of tree-like structure present in the genotypes found for each locus in the complete set of 18 isolates (Hudson, 1998). The sequence alignments were converted to NEXUS files and the split decomposition was performed with SPLITSTREE 2.0.

3. RESULTS AND DISCUSSION

3.1. Discriminatory power of the MLST scheme proposed among wine S. cerevisiae *strains*

Several methods have been proposed for *S. cerevisiae* differentiation at the strain level, as described above. However be found worth trying other techniques that have shown to be useful in alternative organisms for molecular typing of *S. cerevisiae* wine yeast strains.

The differentiation of *S. cerevisiae* at the strain level becomes a major concern, since their adaptation to wine and influence on organoleptic quality are strain specific. Among molecular techniques, restriction analysis of the mitochondrial genome has been successfully used to differentiate between wine yeast strains (Querol et al., 1992). The wine *S. cerevisiae* strains recovered from almost thirty years from various Spanish geographical regions are expected to be diverse. The eighteen strains analyzed in this work showed seventeen different mitochondrial RFLP patterns (Fig. 1). Therefore, the results obtained in this study corroborate that mitochondrial DNA RFLP analysis showed a high discrimination power, with only two strains sharing the same restriction pattern (Fig. 1, Table 1).

One of the main objectives of this study was to investigate the usefulness of MLST to differentiate *S. cerevisiae* strains of enological origin, and additionally, to investigate global genomic similarity at the intraspecific level. The first step in order to perform a MLST analysis of a given species is the adequate choice of the targeted housekeeping genes, the region for amplification and the primer sequences. In this work, several housekeeping genes were initially considered for MLST analysis of wild isolates of *S. cerevisiae*, including the following nuclear

genes: *ALG8*, *CHS1*, *DMC1*, *EGT2*, *EHT1*, *LAP3*, *LYS5*, *MAE1*, *MDH1*, *NOC2*, *RAD57*, *URA3*, *YPS1*, and *YPS3* (data not shown). The sequence of these *S. cerevisiae* genes were aligned with those of the orthologous genes from *Candida glabrata*, *Kluyveromyces lactis*, and *Debaryomyces hansenii*, in order to identify variable regions flanked by conserved regions. These conserved regions would be the target sequences for PCR primers, in order to minimize the risk of unsuccessful amplifications due to strain-to-strain variability among *S. cerevisiae*. As a result of this analysis, primers were designed for the amplification of 400-600 bp regions of the *S. cerevisiae* genes *EHT1*, *LYS5*, *RAD57*, *YPS1* and *YPS3*. These regions were PCR amplified from strains BY4741, EC1118, PMA and IFI1685 in order to ascertain the degree of genetic variability that could be revealed by these loci. The results were disappointing, since the number of SNPs in the subset of strains analyzed varied from 0 to 3, depending on the specific locus (data not shown). In addition to these nuclear genes, four mitochondrial genes were also analyzed, avoiding amplicons spanning intron insertion regions. Again, very little genetic variability was detected (data not shown). Hence we decided to perform the MLST analysis using the loci previously used for the characterization of clinical isolates of *Candida albicans* (Bougnoux et al., 2002), namely *ACC1*, *ADP1*, *ALA1*, *GLN4*, *RPN2* and *VPS13*, respectively orthologs of Ca*ACC1*, Ca*ADP1*, *CaSYA1*, Ca*GLN4*, Ca*RPN2* and Ca*VPS13* from *C. albicans*. Genomic sequences of *S. cerevisiae* and *C. albicans* were aligned and primers for *S. cerevisiae* were designed in equivalent regions (Table 2). In spite of the resulting primers did not meet most of the standard design criteria for PCR primers, they allowed the amplification of the target regions, and the yield obtained was high enough for

automatic sequencing. The single exception was *VPS13*, it was necessary to design new, improved primers in order to be able to amplify the cognate fragment from *S. cerevisiae*, however the success of the amplification was still unreliable and this locus was finally not considered for further analysis. The remaining five loci were used for MLST analysis of the strains listed in table 1. These include commercial wine yeast strains, the laboratory strain BY4741, and *S. cerevisiae* isolates from our historical wine yeast collection. As can be seen in Table 1, these are isolates from diverse Spanish wine producing regions, over the period 1953-1982. Strains from this collection are interesting because they were isolated from almost all wine producing regions of the country, most of them before the use of commercial active dry yeasts was introduced in every specific region, and always before it was introduced in the particular vineyard and cellar. So they are expected to reveal genetic diversity before any potential detrimental effect on it due to new oenological practices would happen.

Only ten polymorphic sites for wine *S. cerevisiae* yeast strains were revealed in the gene fragments analyzed in this study, six for *ALA1*, two for *GLN4*, one each for *RPN2* and *ACC1*, and none for *ADP1*. This resulted in 7, 5, 3 and 2 different genotypes, respectively, for each of these genes (Fig. 2). All the polymorphic sites were bi-allelic, as previously reported for *S. cerevisiae* in studies of yeast biodiversity and phylogeny by SNP analysis for all the polymorphic sites analyzed (Ben-Ari, et al., 2005); 98% of the polymorphic positions detected by Aa et al., (2006); or 97% of those detected by Ayoub et al., (2006). Most strains we analyzed (66%) were homozygous for all the polymorphic sites, and most of the rest were

heterozygous for just one position, but heterozygosis was found in at least one strain for all but two of the positions. The commercial strain EC1118 concentrated most of the heterozygosis found in this study; since half of the positions were heterozygous, suggesting this strain to be the result of a recent (in evolutionary terms) process of hybridization. Finally, most codon changes (6 out of 10) were synonymous, and the rest gave rise to conservative changes.

The combination of the different alleles for each of these genes allowed classifying the 18 strains analyzed in this work in 13 different DST (Table 1). Most DST were strain specific, but there were five strains with a common DST, DST 4, and a pair of strains sharing DST 10. The two strains showing DST 10 were isolated from the velum in Sherry wine aging, in two different cellars, but there was no apparent relationship between the geographical or substrate origin of strains sharing DST 4, since this group included strains from Toro and Ribera de Duero, rather cold regions, as well as La Mancha, a hot and dry region.

Curiously, data obtained from ITS-RFLP and MLST analysis could be complementary, as strains that shared a similar ITS-RFLP pattern (IFI480 and IFI664 strains) could be differentiated by their MLST genotypes. Differences between the respective mitochondrial and nuclear DNA topologies have been also observed by applying MLST analysis to higher eukaryotic organisms (e.g., in the desert night lizard *Xantusia vigilis* species complex) (Leavitt et al., 2007).

The discrimination obtained by MLST in our study of wine *S. cerevisiae* strains was higher than the previously reported by Ben-Ari et al. (2005) who included 14 wine strains in their analysis (about half of the strain in their study), and found them to have almost identical phenotypes and homozygous in most positions. Aa et al. (2006), also included wine yeast strains in their study, from 8 different geographical origins, and analyzed genes (promoter and coding region) which expression might be relevant in wine making (*FZF1, SSU1*), in this case heterozygous positions were not found at all, probably due to the high selective pressure on these loci in winemaking conditions, or sample size. The study by Ayoub et al. (2006), included among other target genes *RPN2*, and found 5 polymorphic positions in this gene, in contrast to the single polymorphic region found in the present work. However, there is not real disagreement with our present results since the amplified region is not overlapping between both studies.

Therefore, it could be concluded from these results that the selection of the gene and the amplified gene fragment are important factors to determine the usefulness of a MLST typing scheme. MLST schemes are based on sequences of multiple (usually seven) loci because the analysis of a single gene provides too little discrimination to be used for molecular typing. As a first step for developing a typing method, we analyzed the sequence diversity of five genes in order to know their usefulness typing discrimination, since the number of loci can be increased to improve resolution, but there will come a point when, for epidemiological purposes, little additional information is obtained for the cost and effort involved (Urwin and Maiden, 2003). In this work, we found only four polymorphic loci, which provide 13

DST among the 18 strains analyzed. Therefore, increased sampling and the use of additional more sensitive genes are needed to establish a highly discriminatory MLST typing straegy for wine *S. cerevisiae* strains.

3.2. Wine S. cerevisiae *population structure*

Mitochondrial DNA has been the workhorse of research in phylogeography of higher eukaryotic organisms for almost two decades. However, concerns with basing evolutionary interpretations on mitochondrial DNA results alone have been voiced since the inception of such studies. Recently, some authors have suggested that species limits are unwarranted unless corroborated by other evidences, usually in the form of nuclear gene data (Zink and Barrowclough, 2008).

An advantage of MLST analysis is that they allowed inferring phylogenetic relationships among the analyzed strains. Concatenated *ADP1*, *ACC1*, *RPN2*, *GLN4*, and *ALA1* gene sequence fragments were analyzed. A cladogram showing the genetic relatedness among the wine *S. cerevisiae* strains investigated in this study is shown in Figure 3. This cladogram, apart of grouping together all the strains having DST 4, which is not obviously revealed by mitochondrial RFLP analysis, suggests a closer relatedness among these strains than between any of them and strains outside that group. In general, pairing between strains in the cladogram does not correspond to any striking similarity between the mitochondrial DNA restriction patterns of these strains; some examples are PMA/IFI466, IFI480/IFI1685, IFI285/IFI664 or IFI475/IFI665. There are however at least two examples of strikingly similar mitochondrial DNA restriction patterns that do not

296 correlate with similarities in the genotype revealed by MLST analysis, these
297 examples are EC1118/PMA and more strikingly IFI480/IFI664/IFI691. This could be
298 some of the potential problems mentioned above, when mitochondrial DNA data
299 are used to infer population structure (Zink and Barrowclough, 2008)
300
301 Examination of the sequences of the analyzed genes can prove evidence for the
302 significance of recombination. Recombination could be detected by the
303 appearance of a network of relationships among sequences rather than a
304 bifurcating tree-like phylogeny. Concatenated gene sequence fragments were
305 analyzed and graphically displayed with SplitsTree (Fig. 4). The split decomposition
306 analysis of the wine *S. cerevisiae* strains analyzed in this work reveals three
307 uncentered edges, suggesting that the evolution of these strains stems from a tree
308 of strains from which single branches radiate. The length of the branches is short,
309 indicating a close relation between the strains analyzed. The relationships among
310 the members of the group were assessed by examining the number of nodes
311 between two isolates.
312
313 Figure 5 shows the split graphs for all the genotypes of the *RPN2*, *GLN4* and *ALA1*
314 polymorphic genes analyzed. The split graph of the *ACC1* gene displays a line
315 because only two genotypes were identified (data not shown). A parallelogram will
316 appear whenever recombination has been involved in the evolution of the analyzed
317 gene. The split graph obtained with *RPN2* and *GLN4* loci showed no evidence of
318 recombinational evolution. We observed parallelogram only in one of the genes
319 analyzed. The *ALA1* locus presents this structure indicating the presence of

homoplasies, probably evolved by intergenic recombination. A point mutation will generate a single nucleotide difference, whereas a recombinatorial exchange is likely to introduce multiple nucleotide differences. The differences in structure among the split graphs obtained for the five loci can be explained by recombination, because recombination can lead to the assembly of genes with different histories within one strain. The conclusion on recombination playing a role in genomic evolution of wine yeast strains is in agreement with that of Puig *et al*. (2000) and Perez-Ortin *et al*. (2002).

The utility of MLST for the analysis of the genetic structure of bacterial pathogens is mainly based on the characteristic of housekeeping genes to have selectively neutral variability. The split decomposition analysis provides evidence that intraspecies recombination occurs in wine *S. cerevisiae* strains and plays a role in generating genetic heterogeneity among strains. The extension of the present analysis to a larger number of isolates could contribute to improved knowledge about the structure of *S. cerevisiae* populations. The relative genetic homogeneity of *S. cerevisiae* wild type strains found in this work is in agreement with previous findings by other authors as described above, and in contrast with the genetic variability encountered in others yeast, like *C. albicans*.

In conclusion, for the purpose of *S. cerevisiae* strain differentiation, mitochondrial RFLP analysis outperforms the MLST analysis described in this work, both in terms of discrimination power and because of its simplicity and lower cost. However, MLST analysis offers the possibility of studying genetic relatedness between yeast

isolates which would be much cumbersome by using mitochondrial RFLP profiles. This study constitutes the first step for the development of an MLST method for wine *S. cerevisiae* strains. Additional discriminating genes will be needed to establish a highly discriminatory MLST typing method for these strains.

Acknowledgements

We are grateful to José Mª Barcenilla and Maria Victoria Santamaría, for their technical assistance, and to Amparo Querol for help in ITS pattern identification. This work was supported by the Spanish Ministerio de Educación y Ciencia (Consolider CSD2007-00063); and the Comunidad de Madrid local Government (grant S-505/AGR-0153).

References

Aa, E., Townsend, J., P., Adams, R. I., Nielsen, K., M., Taylor, J. W., 2006. Population structure and gene evolution in *Saccharomyces cerevisiae*. FEMS Yeast Research 6, pp. 702-715.

Ayoub, M. –J., Legras, J. –L., Saliba, R., Gaillardin, C., 2006. Application of multi locus sequence typing to the analysis of the biodiversity of indigenous *Saccharomyces cerevisiae* wine yeasts from Lebanon. Journal of Applied Microbiology 100, 699-711.

366 Ben-Ari, G., Zenvirth, D., Sherman, A., Simchen, G., Lavi, U., Hillel, J., 2005
367 Application of SNPs for assessing biodiversity and phylogeny among yeast
368 strains. Heredity 95, 493-501.
369 Bougnoux, M. –E., Morand, S., d´Enfert, C., 2002. Usefulness of multilocus
370 sequence typing for characterization of clinical isolates of *Candida albicans*.
371 Journal of Clinical Microbiology 40, 1290-1297.
372 Balerias Couto, M. M., Ejisma, B., Hosftra, H., Huis in´tVelt, J. H. And Vossen, J.
373 M. B. M., 1996. Evaluation of molecular typing techniques to assign genetic
374 diversity among Saccharomyces cerevisiae strains. Applied and
375 Environmental Microbiology 62, 41-46.
376 De las Rivas, B., Marcobal., A., Muñoz, R., 2004. Allelic diversity and population
377 structure in *Oenococcus oeni* as determined from sequence analysis of
378 housekeeping genes. Applied and Environmental Microbiology 70, 7210-
379 7219.
380 De las Rivas, B., Marcobal., A., Muñoz, R., 2006. Development of a multilocus
381 sequence typing method for analysis of *Lactobacillus plantarum* strains.
382 Microbiology 152, 85-93.
383 Dingle, K. E., Colles, F. M., Wareing, D. R. A., Ure, R., Fox, A. J., Bolton, F. E.,
384 Bootsma, H. J., Willems, R. J. L., Urwin, R., Maiden, M. C. J., 2001.
385 Multilocus sequence typing system for *Campylobacter jejuni*. Journal of
386 Clinical Microbiology 39, 14-23.
387 Enright, M. C., Spratt, B. G., 1999. Multilocus sequence typing. Trends in
388 Microbioloby 7, 482-487.

389 Escot, S., Feuillat, M., Dulau, L. and Charpetier C., 2001 Release of
390 polysaccharides by yeasts and the influence of release polysaccharides on
391 color stability and wine astringency, Australian Journal of Grape and Wine
392 Research 7 (3), 153–159.
393 Farfán, M., Miñana-Galbis, D., Fusté, M. C., Lorén, J. G., 2002. Allelic diversity and
394 population structure in *Vibrio cholerae* O139 Bengal based on nucleotide
395 sequence analysis. Journal of Bacteriology 184, 1304-1313.
396 Fay, J. C., Benavides, J. A., 2005. Evidence for domesticated and wild populations
397 of *Saccharomyces cerevisiae*. PLOS Genetics 1, 66-71.
398 Fernández-Espinar, M. T., Esteve-Zardoso, B., Querol, A., Barrio, E., 2000. RFLP
399 analysis of the ribosomal internal transcribed spacers and the 5.8S rRNA
400 gene region of the genus *Saccharomyces*: a fast method for species
401 identification and the differentiation of flor yeasts. Antonie van Leeuwenhoek
402 78, 87-97.
403 Fernández-Espinar, M. T., López, V., Ramón, D., Bartra, E., Querol, A., 2001.
404 Study of the authenticity of commercial wine yeast strains by molecular
405 techniques. International Journal of Food Microbiology 70, 1-10.
406 Fleet G.H., 1993. The microorganisms of winemaking—isolation, enumeration and
407 identification. In: Fleet G.H. (ed.). Wine Microbiology and Biotechnology.
408 Sydney, Australia: Harwood. pp. 2-25.
409 Gallego F. J., Perez M. A., Martinez I., and Hidalgo P., 1998. Microsatellites
410 Obtained From Database Sequences Are Useful to Characterize
411 Saccharomyces cerevisiae Strains. American Journal of Enology and
412 Viticulture, 49 (3), 350-351

Grando, M. S., Ubeda, J. and Briones, A. I., 1994. RAPD analysis of wine Saccharomyces cerevisiae strains differentiated by pulsed field gel electrophoresis. Biotechnology techniques 8 (8), 557-560

Guillamón, J. M., Sabaté, J., Barrio, E., Cano, J., Querol, A., 1998. Rapid identification of wine yeast species based on RFLP analysis of the ribosomal internal transcribed spacer (ITS) region. Archives in Microbiology 169, 387-392.

Helgason, E., Tourasse, N. J., Meisal, R., Caugant, D. A. and Kolsto A. B., 2004. Multilocus Sequence Typing Scheme for Bacteria of the Bacillus cereus Group. Applied and Environmental Microbiology 70 (1), 191-201

Hudson, D. H., 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14, 68-73.

Kunkee R. E. and Amerine M. A., 1970. Yeast in winemaking. In AH Rose, JS Harrison (Eds), The Yeast III, Yeast Technology, Academic Press, London, pp 5-71.

Lavallée, F., Salvas, Y., Lamy, S., Thomas, D. Y., Degré, R. and Dulau, L., 1994. PCR and DNA fingerprinting used as quality control in the production of wine yeast strains. American Journal of Enology and Viticulture 45, 86-91

Leavitt, D. H., L., Bezy, R. L., Crandall K. A., Sites, J. W. Jr., 2007. Multi-locus DNA sequence data revealed a history of deep crytic vicariance and habitat-driven convergence in the desert night lizard *Xantusia vigilis* species complex (Squamata: Xantusiidae). Molecular Ecology 16, 4455-4481.

435 Legras, J. L. and Karst, F., 2003. Optimisation of interdelta analysis for
436 Saccharomyce cerevisiaestrains characterization. FEMS Microbiology
437 Letters 221, 249–255.
438 López, V., Fernández-Espinar M. T., Barrio, E., Ramón, D., Querol, A. 2003. A new
439 PCR-based method for monitoring inoculated wine fermentations.
440 International Journal of Food Microbiology 81, 63-71.
441 Maiden, M. C., Bygraves, J. J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R.,
442 Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M.,
443 Spratt, B. G., 1998. Multilocus sequence typing: a portable approach to the
444 identification of clones within populations of pathogenic microorganisms.
445 Proceedings of the National academy of Sciences USA 95, 3140-3145.
446 Martínez-Rodriguez, A., Carrascosa, A. and Polo, M.C., 2001. Release of nitrogen
447 compounds to the extracellular médium by three strains of *Saccharomyces*
448 *cerevisiae* during induced autolysis in a model wine system. International
449 Journal of Food Microbiology 68, 155-160.
450 Hay más articules de Martinez-Rodriguez et al. 2001, compruebalo
451 Ness, F., Lavallee, F., Dubourdieu, D., Aigle, M. and Dualu, L., 1993. Identification
452 of yeast strains using the polymerase chain reaction. Journal of the Science
453 of Food and Agriculture 62, 89-94.
454 Pérez-Ortín, J. E., Querol, A., Puig, S., Barrio, E., 2002. Molecular characterization
455 of a chromosomal rearrangement envolved in the adaptative evolution of
456 yeast strains. Genome Research 12, 1533-1539.

457 Puig, S., Querol, A., Barrio, E., Pérz-Ortín, J. E., 2000. Mitotic recombination and
458 genetic changes in *Saccharomyces cerevisiae* during wine fermentation.
459 Applied and Environmental Microbiology 66, 2057-2061.
460 Querol A. and Fleet G. (eds), 2006. Yeast in food and Beverages (The yeast
461 Handbook; 2) Springer-Verlag Heidelberg
462 Querol, A. and Ramon, D., 1996. The application of molecular techniques in wine
463 microbiology. Trends in Food Science and Technology 7, 73-78.
464 Querol, A., Barrio, E., Huerta, T., Ramón, D., 1992. Molecular monitoring of wine
465 fermentations conducted by dry yeast strains. Applied and Environmental
466 Microbiology 58, 2948–2953.
467 Quesada, M. P. and Cenis, J. L., 1995. Use of Random Amplified Polymorphic
468 DNA (RAPD-PCR) in the Characterization of Wine Yeasts. American
469 Journal of Enology and Viticulture 46 (2), 204-208.
470 Robles J. C., Koreen L., Park, S., Perlin D. S., 2004. Multilocus sequence typing is
471 a reliable alternative method to DNA fingerprinting for discriminating among
472 strains of *Candida albicans*. Journal of Clinical Microbiology 42, 2480-2488.
473 Sullivan C. B., Diggle M. A. and Clarke S. C. 2005. Multilocus sequence typing:
474 Data analysis in clinical microbiology and public health. Molecular
475 Biotechnology 29 (3), 245-54.
476 Techera, A. G., Jubany, S., Carrau, F. M. and Gaggero, C. 2001. Differentiation of
477 industrial wine yeast strains using microsatellite markers. Letters in Applied
478 Microbiology 33, 71–75.
479 Torija, M. J., Rozès, N., Poblet, M., Guillamón, J. M. and Mas, A., 2001. Yeast
480 population dynamics in spontaneous fermentations: comparison between

two different wine areas during three consecutive years. Antonie van Leeuwenhoek 79, 345–352.

Urwin, R., Maiden M. C. J., 2003. Multi-locus sequence typing: a tool for global epidemiology. TRENDS in Microbiology 11, 479-487.

Vezinhet, F., Blondin, B. and Hallet, J. N., 1990. Chromosomal DNA patterns and mitochondrial DNA polymorphism as tools for identification of enological strains of *Saccharomyces cerevisiae*. Applied Microbiology and Biotechnology 32, 568–571.

Zink, R. M., Barrowclough, G. F., 2008. Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology 17, 2107-2121.

Figure captions

Figure 1.Mitochondrial RFLP analysis of the 18 oenological *S. cerevisiae* strains examined in this study. Chromosomal DNA was digested with the restriction enzyme *Hin*fI. The molecular marker *Eco*RI + *Hin*dIII digested λ DNA is showed in the first and last lines of the figure.

Figure 2. Polymorphic sites in each of the four genes fragments analyzed. Each of the sites where the sequence of one or more of the genes differs is shown; only sites that differ are shown, sites that are identical are indicated by periods. The number od strains possessing the allele is indicated in parenthesis. Numbering on the polymorphic sites (vertical format) is from the first nucleotide position of the corresponding gene (Y = T or C, R = G or A, and W = A or T).

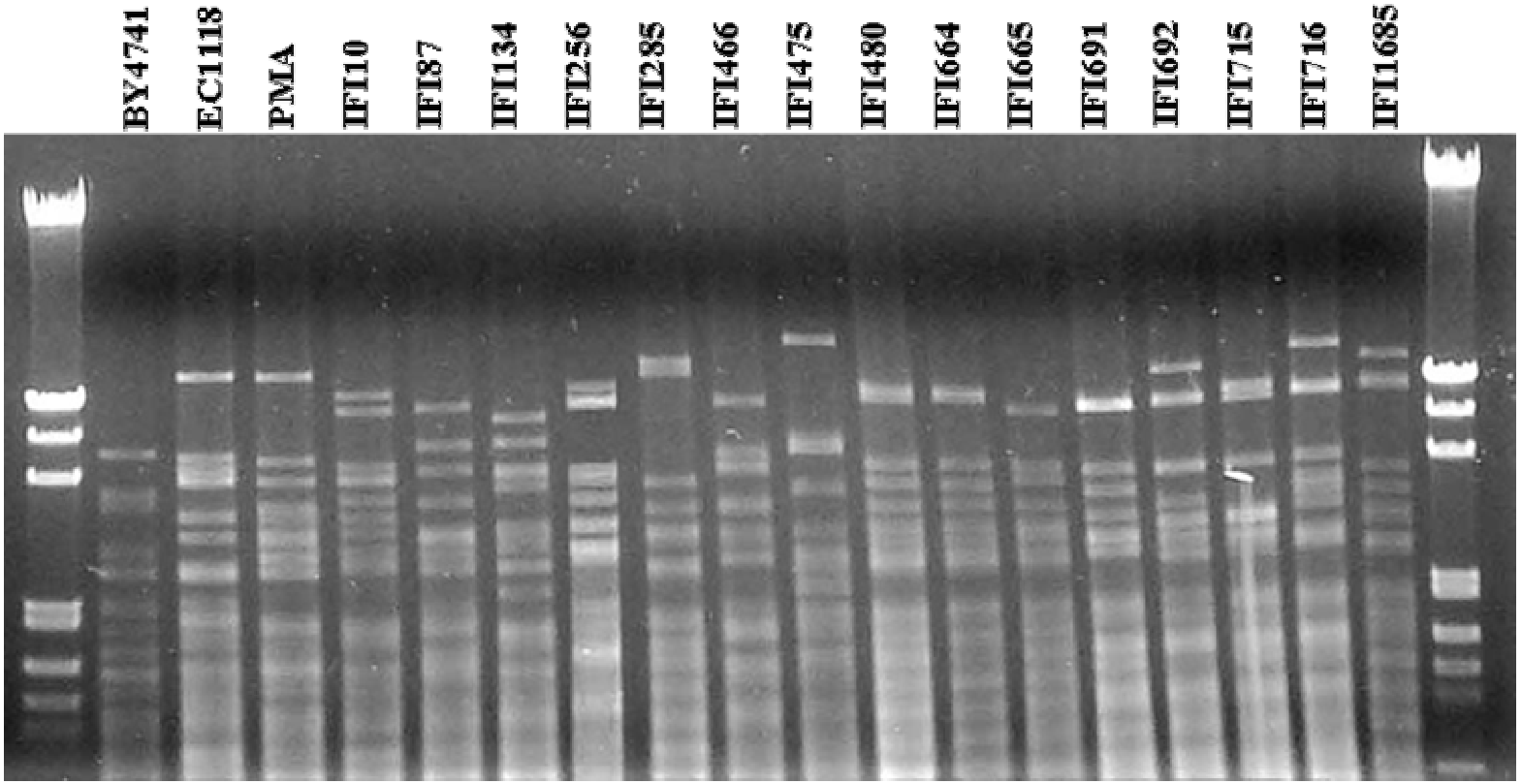
Figure 3. Cladogram showing the genetic relatedness of the 18 wine *S. cerevisiae* strains examined in this study. The cladogram was constructed from the sequence of concatenated *ADP1*, *ACC1*, *RPN2*, *GLN4*, and *ALA1* gene fragments.

Figure 4. Split decomposition analysis based on the allelic profiles of the 18 wine *S. cerevisiae* examined in this study.

Figure 5. Split decomposition analysis of genotypes obtained from 18 *S. cerevisiae* strains from the locus possessing, at least, three polymorphic positions. The observation that in the *ALA1* graph several alleles in the sample are connected to

each other by multiple pathways, forming parallelograms structures, is suggestive of recombination. All branch lengths are draw to scale. The numbering refers to genotype numbers.

Figure 1


BY4741
EC1118
PMA
IFI10
IFI87
IFI134
IFI256
IFI285
IFI466
IFI475
IFI480
IFI664
IFI665
IFI691
IFI692
IFI715
IFI716
IFI1685

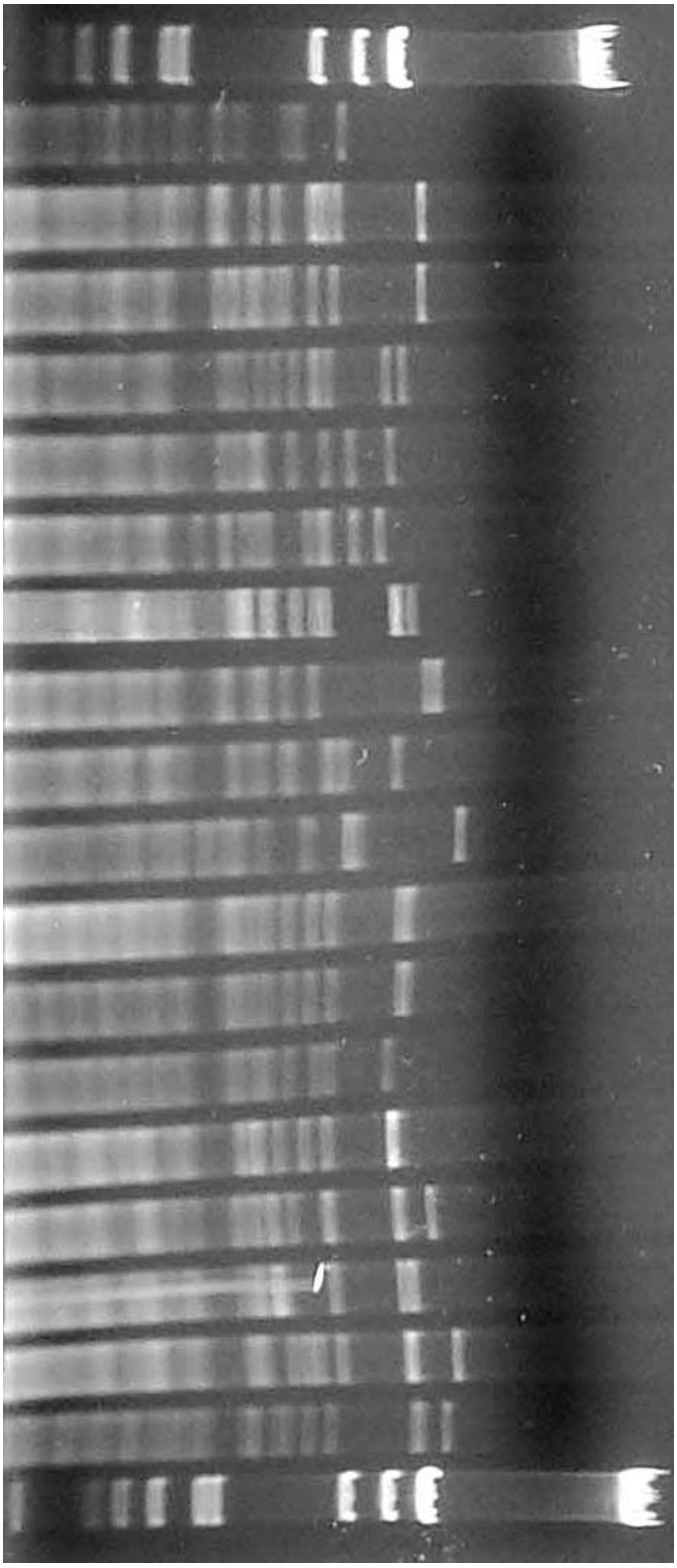

BY4741
EC1118
PMA
IFI10
IFI87
IFI134
IFI256
IFI285
IFI466
IFI475
IFI480
IFI664
IFI665
IFI691
IFI692
IFI715
IFI716
IFI1685

Figure 2

Figure 2

ACC1

	3 4 7 1
genotype 1 (15)	T
genotype 2 (1)	Y

RPN2

	1 2 3 7
genotype 1 (10)	G
genotype 2 (5)	A
genotype 3 (1)	R

GLN4

	1 3 8 0 2 3
genotype 1 (9)	T A
genotype 2 (2)	. G
genotype 3 (3)	. R
genotype 4 (3)	A .
genotype 5 (1)	W A

ALA1

	2 2 2 2 2 2 4 6 6 6 6 6 2 0 0 2 4 6 4 1 7 0 4 4
genotype 1 (7)	A C G G C T
genotype 2 (1)	G . . . T C
genotype 3 (6)	G T . . T C
genotype 4 (1)	. . A . . .
genotype 5 (1)	. . R . . .
genotype 6 (1)	G . . T T .
genotype 7 (1)	R Y . . Y Y

Figure 3

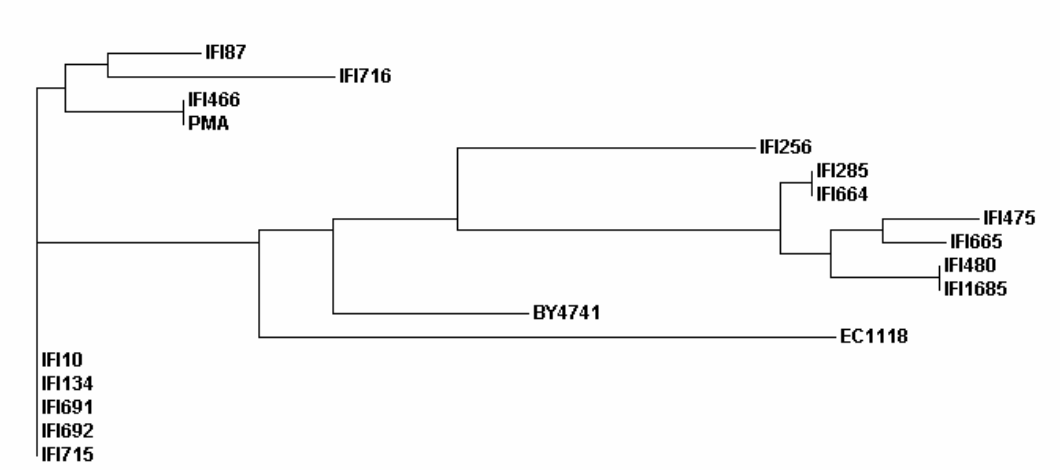

IFI87
IFI716
IFI466
PMA
IFI256
IFI285
IFI664
IFI475
IFI665
IFI480
IFI1685
BY4741
EC1118
IFI10
IFI134
IFI691
IFI692
IFI715

Figure 4

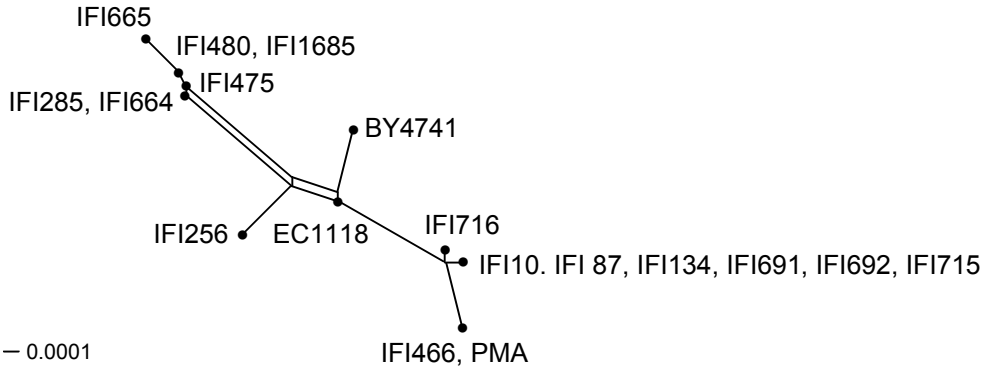

IFI665
IFI480, IFI1685
IFI475
IFI285, IFI664
BY4741
IFI256
EC1118
IFI716
IFI10. IFI 87, IFI134, IFI691, IFI692, IFI715
IFI466, PMA
0.0001

Figure 5

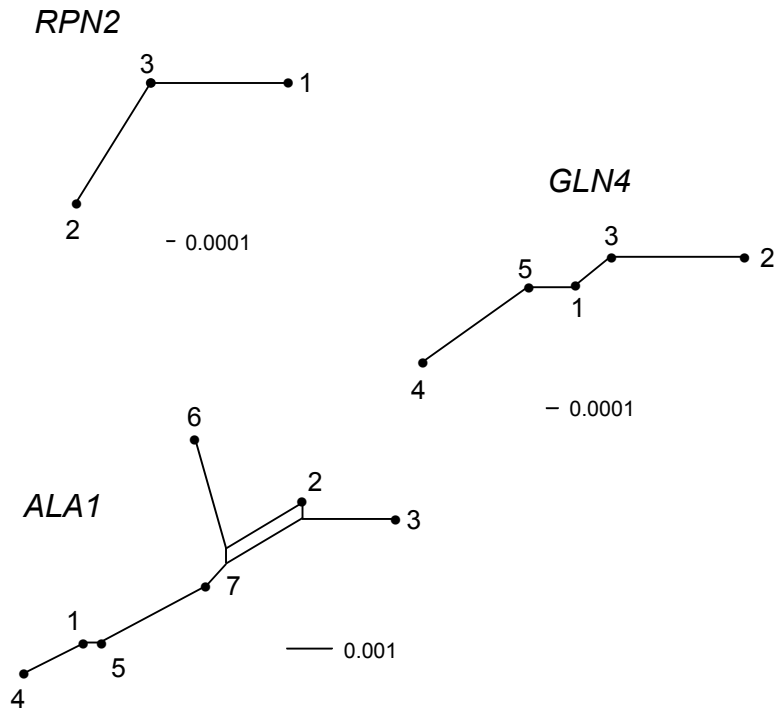


Table 1

Table 1

Properties of wine *S. cerevisiae* isolates analyzed and their allele profile at each locus

Strain	ITS-RFLP[a]	Mt-RFLP[b]	DST[c]	Genotype no. at locus					Source of isolate		
				ADP1	*ACC1*	*RPN2*	*GLN4*	*ALA1*	Substrate	Geog. origin	Year
BY4741	A	1	1	1	1	1	1	6	Laboratory strain		
EC1118	A	2	2	1	1	1	3	7	Commercial strain		
PMA	A	3	3	1	1	1	4	1	Commercial strain		
IFI10	A	4	4	1	1	1	1	1	Grapes (Airén)	La Mancha	1953
IFI87	A	5	5	1	1	1	1	4	Grapes (P.Ximenez)	Montilla	1957
IFI134	A	6	4	1	1	1	1	1	Grapes (Salema)	Huelva	1958
IFI256	A	7	6	1	2	1	4	2	Must (Albariño)	Rias Baixas	1974
IFI285	B	8	7	1	1	2	1	3	Must	Rueda	1975
IFI466	A	9	8	1	1	1	4	1	Grapes (P.Ximenez)	Cordoba	1957
IFI475	A	10	9	1	1	3	2	3	ND[d]	San Sadurni	1958
IFI480	B	11	10	1	1	2	3	3	Wine (“Flor”)	Sevilla	1958
IFI664	B	11	11	1	1	2	1	3	Wine (“Flor”)	Montilla	1958
IFI665	A	12	12	1	1	2	2	3	Grape (Salema)	Huelva	1986
IFI691	A	13	4	1	1	1	1	1	Grape (Sherry)	Cigales	1982
IFI692	A	14	4	1	1	1	1	1	Grape (Malvasia)	Toro	1982
IFI715	A	15	4	1	1	1	1	1	Must	Xxxx,	1982
IFI716	A	16	13	1	1	1	5	5	Must	Málaga	1982
IFI1685	A	17	10	1	1	2	3	3	Wine (“Flor”)	Jerez	ND

[a] ITS-RFLP, type based on the enzyme RFLP analysis of the 5.8S-ITS region
[b] Mt-RFLP, type based on the *Hin*fI mitochondrial DNA extraction restriction pattern
[c] DST, Diploid sequence type
[d] ND, Data not available

Table 2
Primers used for MLST typing scheme of wine *S. cerevisiae* strains

ORF	Gene	Primers	Sequence 5′→3′	5′start position[a]	PCR product length (bp)
YNR016C	*ACC1*	ACC1F	GCAAGAGAAATTTTGATTCAAGG	3073	492
		ACC1R	TTCATCAACATCATCTAAATG	3564	
YKK040C	*VPS13*	2VPS13F	ATTTCACTTAGAGATATTCGTCT GGC	4981	830
		2VPS13R	TTTTGCCCAGAAACACAAACACC	5810	
YOR168W	*GLN4*	GLN4F	GAGATTGTCAAGAATAAAAAGG T	67	489
		GLN4R	GTCTCTCTCATCCTTTGGACC	555	
YCR011C	*ADP1*	ADP1F	GAGCCTTCTATGAATGATTTG	826	585
		ADP1R	TTGATCGACGAACCCGATTAT	1410	
YIL075C	*RPN2*	RPN2F	TTTATGCACGCTGGTACTAC	1012	450
		RPN2R	GAGACCCATACCTAATGCAG	1461	
YOR335C	*ALA1*	ALA1F	AGAAGAATTGTTGCTGTTACTG	2236	552
		ALA1R	ATTACCTTTACCACCAGCCTT	2787	

[a] Being position 1 the first of the ATG start codon