ORIGINAL PAPER

The genetic structure of fermentative vineyard-associated *Saccharomyces cerevisiae* populations revealed by microsatellite analysis

Dorit Schuller · Margarida Casal

Abstract From the analysis of six polymorphic microsatellite loci performed in 361 *Saccharomyces cerevisiae* isolates, 93 alleles were identified, 52 of them being described for the first time. All these isolates have a distinct mtDNA RFLP pattern. They are derived from a pool of 1620 isolates obtained from spontaneous fermentations of grapes collected in three vineyards of the Vinho Verde Region in Portugal, during the 2001–2003 harvest seasons. For all loci analyzed, observed heterozygosity was 3–4 times lower than the expected value supposing a Hardy–Weinberg equilibrium (random mating and no evolutionary mechanisms acting), indicating a clonal structure and strong populational substructuring. Genetic differences among *S. cerevisiae* populations were apparent mainly from gradations in allele frequencies rather than from distinctive "diagnostic" genotypes, and the accumulation of small allele-frequency differences across six loci allowed the identification of population structures. Genetic differentiation in the same vineyard in consecutive years was of the same order of magnitude as the differences verified among the different vineyards. Correlation of genetic differentiation with the distance between sampling points within a vineyard suggested a pattern of isolation-by-distance, where genetic divergence in a vineyard increased with size. The continuous use of commercial yeasts has a limited influence on the autochthonous fermentative yeast population collected from grapes and may just slightly change populational structures of strains isolated from sites very close to the winery where they have been used. The present work is the first large-scale approach using microsatellite typing allowing a very fine resolution of indigenous *S. cerevisiae* populations isolated from vineyards.

Keywords Microsatellite · *S. cerevisiae* · Commercial yeasts · Yeast population

Introduction

The initial stages of traditional spontaneous wine fermentations are carried out by yeast species that are present on the grape's surface such as the apiculate yeasts *Hanseniaspora uvarum* (= *Kloeckera apiculata*) and other yeasts belonging to the genera *Metschnikowia, Candida* or *Pichia*, together with moulds, lactic and acetic acid bacteria (Fleet and Heard 1993). Contrarily, *Saccharomyces cerevisiae*, the predominant yeast species used in the production of wine, universally known as "wine yeast", occurs in extremely low

D. Schuller (✉) · M. Casal
Centro de Biologia (CB-UM), Departamento de Biologia, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal
e-mail: dschuller@bio.uminho.pt

number on healthy undamaged berries or in soils (Frezier and Dubourdieu 1992; Martini et al. 1996; Parish and Carroll 1985), while damaged grapes are believed to be an important source of this species (Mortimer and Polsinelli 1999). The grape's yeast flora depends on a variety of factors such as climatic conditions including temperature and rainfalls, geographic localization of the vineyard (Longo et al. 1991; Parish and Carroll 1985), antifungal applications (Monteil et al. 1986), grape variety, the vineyard's age (Martini et al. 1980; Pretorius et al. 1999; Rosini 1982), as well as the soil type (Farris et al. 1990).

Under the selective conditions of grape must fermentation and with increasing concentrations of ethanol, yeast species of the early fermentative stages are rapidly outgrown by *S. cerevisiae* and related species, which dominate the later stages of the process. The prevalence of *S. cerevisiae* strains is well documented among the wineries resident flora (Beltran et al. 2002; Constanti et al. 1997; Longo et al. 1991; Sabate et al. 2002; Vaughan-Martini and Martini 1995).

Autochthonous *S. cerevisiae* strains isolated from natural environments associated with the wine production areas of interest, obtained from clonal selection, are now commercialized as active dry yeast. Such strains are able to efficiently ferment grape musts and produce desirable metabolites (e.g. glycerol, organic acids and higher alcohols), associated with reduced off-flavors development (mainly H_2S, acetic acid or phenolic compounds). Globally, they enhance the wine's sensorial characteristics and confer typical attributes to specific wine styles (Briones et al. 1995; Regodon et al. 1997). About 200 *S. cerevisiae* wine strains are currently available and their specific application is recommended according to the wine style and/or grape variety. Commercially available yeast starters are now widely used in winemaking without any special containment and are annually released in large quantities, together with liquid and solid wine-making residues, in the environment around the winery. From an ecological point of view, these yeasts can be regarded as non-indigenous strains that are every year introduced in large quantities in the ecosystem surrounding a winery. In a recent study that was carried out in six vineyards of the Vinho Verde

(Portugal) and the Languedoc (France) wine regions, it was shown that the dissemination of commercial yeast strains is limited to a very close proximity of the winery (10–200 m) where they have been used. They were mostly found in samples collected after the onset of wine production, indicating immediate dissemination and their presence in the vineyard was restricted to short distances and limited periods of times showing natural fluctuations of periodical appearance/disappearance like autochthonous strains. Their permanent implantation in the vineyard did not seem to occur (Valero et al. 2005).

The genetic diversity of autochthonous *S. cerevisiae* strains from wine-producing regions has been analyzed by molecular methods such as karyotyping by pulse field gel electrophoresis (Blondin and Vezinhet 1988), mitochondrial DNA restriction analysis (mtDNA RFLP) (Querol et al. 1992) and fingerprinting based on repetitive delta sequences (Legras and Karst 2003; Ness et al. 1993). The most recent molecular technique that is able to resolve this diversity is based on repetitive microsatellite sequences, which are tandem motifs from 1 to 6 bases. Recently, an increasing number of microsatellites have been described for *S. cerevisiae*, with the aim finding the most polymorphic loci with a high allelic diversity that can be applied for both strain delimitation and the description of relationships between strains that are related due to their common geographical or technological origin (Bradbury et al. 2006; Gallego et al. 1998; Hennequin et al. 2001; Legras et al. 2005; Pérez et al. 2001). It has been previously shown that the discriminatory power of six microsatellite loci (Pérez et al. 2001) is identical both to the mtDNA RFLP (using enzyme *Hinf*I) and the optimized interdelta sequence method (Schuller et al. 2004).

With the aim of gaining insight in the genetic variability and populational structure of fermentative vineyard-associated *S. cerevisiae* populations, in the present work the analysis of six polymorphic microsatellite loci was performed in 361 *S. cerevisiae* isolates, previously screened by mtDNA RFLP from a pool of 1620 isolates. All isolates were obtained from spontaneous fermentations of grapes collected in three vineyards

of the Vinho Verde Region in Portugal, during the 2001–2003 harvest seasons. We also evaluated the effect of commercial yeast strains on the yeast populations found in vines surrounding the wineries where such strains are continuously used.
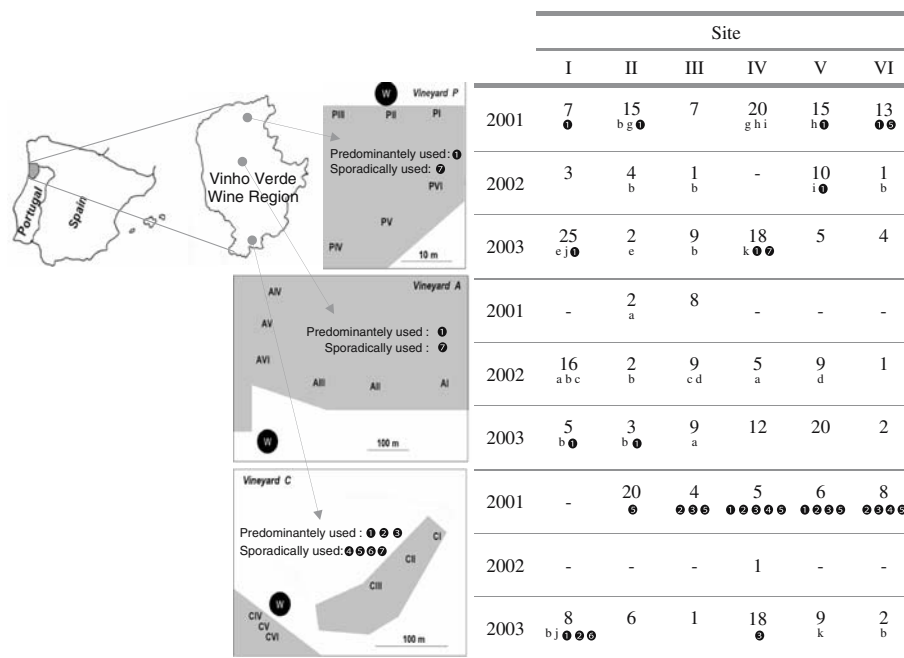
Materials and methods

Sampling

The sampling plan included a total of 18 sites in three vineyards surrounding a winery, located in northwest Portugal (Região Demarcada dos Vinhos Verdes), as shown in Fig. 1. In each vineyard, six sampling points were defined, located at 10–400 m from each other, according to the vineyard geography. In three consecutive years (2001–2003), duplicate grape samples were collected, a few days before and after harvest, respectively, although the grapes were not always collected from the same rootstock, but from the same area

(± 1–2 m). The grapevine varieties sampled were Loureiro (vineyard A), Alvarinho (vineyard P) and Avesso (vineyard C), being all white grapes cultivated in the Vinho Verde Region.

Fermentation and strain isolation

From each sampling point, approximately 2 kg of grapes were aseptically collected and the extracted grape juice was fermented at 20°C in small volumes (500 ml), with mechanical agitation (20 rpm). Fermentation progress was monitored by daily weight determinations. When must weight was reduced by 70 g/l, corresponding to the consumption of about 2/3 of the sugar content, diluted samples (10^{-4} and 10^{-5}) were spread on YPD plates (yeast extract, 1% w/v; peptone 1% w/v; glucose 2% w/v; agar 2%, w/v), and 30 randomly chosen colonies were collected after incubation (2 days, 28°C). The isolates obtained throughout this work were stored in glycerol (30%, v/v) at −80°C.

		Site				
	I	II	III	IV	V	VI
Vineyard P						
2001	7 ❶	15 b g ❶	7	20 g h i	15 h ❶	13 ❶❺
2002	3	4 b	1 b	-	10 i ❶	1 b
2003	25 e j ❶	2 e	9 b	18 k ❶❷	5	4
Vineyard A						
2001	-	2 a	8	-	-	-
2002	16 a b c	2 b	9 c d	5 a	9 d	1
2003	5 b ❶	3 b ❶	9 a	12	20	2
Vineyard C						
2001	-	20 ❸	4 ❷❸❺	5 ❶❷❸❹❺	6 ❶❷❸❹	8 ❷❸❹❺
2002	-	-	-	1	-	-
2003	8 b j ❶❷❸	6	1	18 ❸	9 k	2 b

Fig. 1 Geographic location of the three vineyards A, C and P in the Vinho Verde Region, with indication of the sampling sites (PI-PVI, AI-AVI and CI-CVI) and the wineries (W). The table summarizes the number of strains with unique genotypes for each sampling site and year. The same superscript letters (a–q) represent identical genotypes in different samples. Genotypes of commercial yeast strains, that were isolated from different samples, are indicated by numbers (❶ Zymaflore VL1; ❷ Zymaflore F10; ❸ Zymaflore F15; ❹ Uvaferm BDX; ❺ ICV D254; ❻ Zymaflore VL3; ❼ Lalvin Cy 3079)

DNA isolation

Yeast cells were cultivated in 1 ml YPD medium (36 h, 28°C, 160 rpm). DNA isolation was performed as described (Lopez et al. 2001) with a modified cell lysis procedure, using 25 U of Zymolase (SIGMA). Cell lysis was dependent on the strain and lasted between 20 min and 1 h (37°C). DNA was used for mitochondrial RFLP and microsatellite analysis.

Mitochondrial DNA restriction patterns

Mitochondrial DNA restriction of all strains was carried out as a first screening approach, to reduce the number of isolates to be analysed by microsatellite typing. Digestion reactions were carried out overnight at 37°C and contained 15µl of the previously isolated DNA, and were prepared as previously described (Schuller et al. 2004), in a final volume of 20µl. To each isolate, a pattern designation was attributed (A1–A92, C1–C70 and P1–P135 for isolates from vineyard A, C and P, respectively). When isolates from different samples showed identical patterns, one representative strain from each sample was randomly withdrawn, resulting in a total of 361 isolates that were further studied by microsatellite analysis.

Microsatellite amplification

The six trinucleotide microsatellite loci described as ScAAT1, ScAAT2, ScAAT3, ScAAT4, ScAAT5 and ScAAT6 (Table 1) (Pérez et al. 2001a, b) were amplified and analyzed as previously described (Schuller et al. 2004).

Computer assisted analysis

Based on the the genome sequence for strain S288C (SGD database, http://www.yeastgenome.org/), and the results obtained for the size of microsatellite amplicons of this strain, the number of repeats for alleles from each locus was calculated. Genetic analysis was performed using the software Arlequin 2000 (Schneider et al. 1997) and included (i) estimation of allelic frequencies (ii) observed heterozygosity compared to expected values, (iii) estimation of Wright's F_{ST}

Table 1 Characteristics of the six microsatellite loci ScAAT1–ScAAT6 that were used as genetic markers in the present study

Microsatellite designation	Repeat	ORF or coordinates	Chromosome	Primers	Fluorochrome	Size (S288C)	No. of repeats (S288C)
ScAAT1	ATT	86,901–87,129	XIII	F: AAAAGCGTAAGCAATGGTGT AGAT R: AGCATGACCTTTACAATTTGATAT	6-FAM	229	35
ScAAT2	ATT	YBL084c	II	F: CAGTCTTATTGCCTTGAACGA R: GTCTCCATCCTCCAAACAGCC	HEX	393	20
ScAAT3	ATT	YDR160w	IV	F: TGGGAGGAGGGAAATGGACAG R: TTCAGTTACCCGCACAATCTA	6-FAM	268	23
ScAAT4	ATT	431,334–431,637	VII	F: TGCGGAAGACTAAGACAATCA R: AACCCCATTTCTCAGTCGGA	TET	304	12
ScAAT5	TAA	897,028–897,259	XVI	F: GCCAAAAAAAATAATAAAAAA R: GGACCTGAACGAAAAGAGTAG	TET	231	13
ScAAT6	TAA	105,661–105,926	IX	F: TTACCCCTCTGAATGAAAACG R: AGGTAGTTTAGGAAGTGAGGC	HEX	266	19

value (Wright 1978) and (iv) genetic variation attributable to different hierarchical levels of defined genetic structures (AMOVA analysis). Wright's F_{ST} value was calculated to determine population differentiation among vineyards, among sampling years and also among sampling locations within a vineyard.

An allelic frequencies matrix was obtained based on Euclidean distance and clustered by the unweighted pair group method arithmetic mean (UPGMA) using the program NTSYSpc 2.0 (Applied Biostatistics Inc.) to examine whether genetic divergence was correlated with sampling sites. This software was also used for dendrogram drawing and to calculate a cophenetic correlation coefficient (r).

Results

Recovery of *S. cerevisiae* strains

As shown in Fig. 1, six sampling sites in each of three vineyards, located in the Vinho Verde Wine Region, were sampled during the 2001–2003 harvest seasons. Two sampling campaigns were performed, one before and another after the harvest, in a time frame of about 2 weeks as an attempt to obtain an elevated number of different strains. A total of 108 grape samples have been planned (6 sampling points × 2 sampling campaigns × 3 vineyards × 3 years), from which 54 started a spontaneous fermentation, 36 were not able to start fermentation after 30 days of incubation, whereas 18 samples were not collected due to unfavorable weather conditions and a bad sanitation state of the grapes in 2002. From the 54 fermentations 1620 yeast isolates were obtained. All the isolates were analyzed by their mtDNA RFLP (*Hinf*I) and a pattern profile was attributed to each isolate, resulting in a total of 297 different profiles. The results of this ecological survey, including the temporal and spatial distribution of the found strains have been recently published (Schuller et al. 2005). When the same profile was found in more than one sample, one strain from each sample was randomly withdrawn resulting in a total of 361 isolates, all assumed to be

S. cerevisiae strains. This was supported by their inability to grow in a medium containing lysine as sole nitrogen source and by their capacity to amplify the previously described *S. cerevisiae* specific microsatellite loci ScAAT1–ScAAT6 (Pérez et al. 2001a, b).

The species *S. cerevisiae* is very closely related to the species *Saccharomyces bayanus*, *Saccharomyces pastorianus*, *Saccharomyces paradoxus*, *Saccharomyces cariocanus*, *Saccharomyces mikatae*, and *Saccharomyces kudriavzevii* (Naumov et al. 2000). These six species, together with *S. cerevisiae*, constitute the *Saccharomyces* sensu stricto complex. Only *S. cerevisiae*, *S. bayanus*, *S. pastorianus*, and *S. paradoxus* are associated with fermentative processes. *S. cerevisiae* and *S. bayanus* are considered the predominating species in wine fermentation. *S. paradoxus* has been isolated only once in wine (Redzepovic et al. 2002), whereas *S. pastorianus* is only present in beer making. Our (unpublished) results showed that the specific microsatellite primers do not amplify the homologous loci from other *Saccharomyces* species such as *S. bayanus* and *S. paradoxus*. Sequence analysis was performed with data obtained from the Washington University Genome Sequencing Center (http://www.genome.wustl.edu/projects/yeast/) and the Broad Institute (http://www.broad.mit.edu/annotation/fungi/comp_yeasts/). Both *S. bayanus* and *S. paradoxus* showed no homology with the ScAAT1, ScAAT3, and ScAAT6 primer binding regions. ScAAT4, ScAAT5 and ScAAT2, ScAAT5 primer binding sites had a low homology with the corresponding sequences in *S. bayanus* and *S. paradoxus*, respectively.

Strains showing different mtDNA RFLP patterns had distinct genotypes as determined by the allelic combinations for loci ScAAT1–ScAAT6. Microsatellite analysis performed in a ramdomly selected group of 50 isolates (among the whole collection comprising 1620 strains) showed that isolates with the same/different microsatellite amplification profiles always corresponded to the same/different mtDNA RFLP patterns. In addition, 90 isolates with identical mtDNA RFLP were analyzed in 6 microsatellite loci and always showed the same allelic combinations (our unpublished results). Therefore, allele

frequencies correspond to a random sampling of the alleles present in the microfermentations.

The table in Fig. 1 indicates the number of different microsatellite genotypes obtained from strains collected at each sampling site in both sampling campaigns (before and after the harvest). The number of different strains isolated from each sampling point showed a lower (1–10 strains) or higher (11–21 strains) biodiversity. Genotypes a–k showed a wider temporal and geographical distribution, being the corresponding strains characterized by a generalized pattern of sporadic presence, absence and reappearance across sampling sites, vineyards or years. Genotype b showed a more regional distribution with a perennial behavior. In several sampling sites commercial strains were recovered, that have been used predominately (in higher quantity and continuously) or sporadically (in lower quantity and not continuously) by the wineries during the harvests preceeding the 5 years of the current study. The respective genotypes are shown in Table 2. A detailed analysis regarding their predominance and spatio-temporal distribution, including also the results from an identical study performed in the Languedoc wine region (France) has been recently published (Valero et al. 2005).

Genetic analysis of alleles obtained for loci ScAAT1–ScAAT6

The distribution of overall and vineyard-specific allelic frequencies for the loci ScAAT1–ScAAT6 is shown in Fig. 2. The six markers revealed a high degree of genetic variability, ScAAT1 and ScAAT3 being the most polymorphic markers with 29 and 19 alleles, respectively. Besides the 41 alleles (51 strains) previously described for ScAAT1–ScAAT6 (Pérez et al. 2001), 52 new alleles were identified in the present study (361 strains). In general, the most frequent alleles have been previously described, and their distribution is similar in the three vineyards A, C and P. However, we identified some alleles, described for the first time in the present study, that show a surprising high allelic frequency (allele 28, ScAAT1; allele 7, ScAAT2; allele 20, ScAAT3) and could be indicative of the *S. cerevisiae* populations from the Vinho Verde Region.

Table 2 Genotypes expressed as number of trinucleotidic repeats for microsatellite markers ScAAT1-ScAAT6 for the commercial yeast strains that were recovered in different sampling sites and that have been used by the wineries during the 2001–2003 sampling campaigns and in previous years

Designation (Fig. 1)	Commercial name, origin	ScAAT1		ScAAT2		ScAAT3		ScAAT4		ScAAT5		ScAAT6	
❶	Zymaflore VL1, Gironde (F)	29	34	12	15	22		20		15	16	16	17
❷	Zymaflore F10, Bordelais (F)	26		14		22		20		16		16	
❸	Zymaflore F15, Gironde (F)	28		14		16		20		16		16	
❹	Uvaferm BDX, Gironde (F)	28		14		14	26	12	20	16		16	
❺	Lalvin ICV D254, Languedoc (F)	26		14		14		20		15		16	
❻	Zymaflore VL3, Gironde (F)	33	34	12	14	16	22	20		14	15	16	17
❼	Lalvin Bourgoblanc Cy3079, Bourgogne (F)	26	32	14	15	15	21	20		15	16	16	
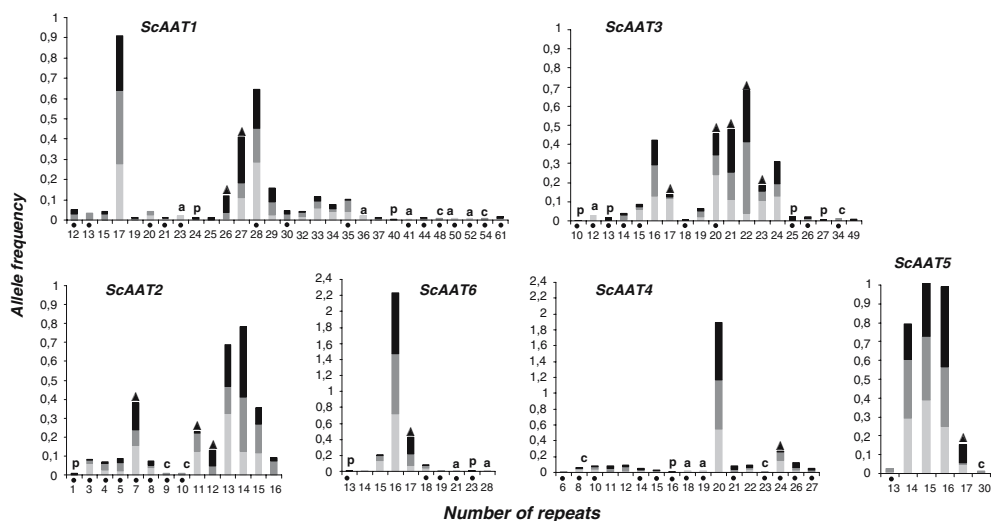
Fig. 2 Alleles of microsatellite loci ScAAT1–ScAAT6 and their frequencies in *S. cerevisiae* in each of the vineyards A (light grey bars), C (dark grey bars) and P (black bars). ● New alleles, identified in the present study; ▲ Alleles with major differences regarding their frequency of occurrence in each vineyard; a, c, p Unique alleles, occurring only in vineyards A, C and P, respectively

Populations from C and P share the most frequent alleles for markers ScAAT1, ScAAT2 and ScAAT3 (17, 14 and 22), while populations belonging to A had the highest frequencies at alleles 28, 13 and 20, respectively. For ScAAT4 and ScAAT6, alleles 20 and 16 were the most frequent for all 3 populations, and for locus ScAAT5 the allele 16 was most frequent in A and C, and allele 15 in P, respectively. Many of the alleles occurring with a lower global frequency, showed different incidences for *S. cerevisiae* populations from vineyards A, C and P (e.g. allele 26 and 27, ScAAT1; allele 7, 11 and 12, ScAAT2; allele 17 and 23, ScAAT3; allele 24, ScAAT4; allele 17, ScAAT5; allele 17; ScAAT6). For each locus, unique alleles were also found in each of the three populations; their frequencies were very low, ranging between 0.01 and 0.03, and they might play only a minor role.

For the populations from different vineyards the observed heterozygosity (Ho) was in general about 3–4 times lower than the expected heterozygosity (He) for all loci analyzed (Table 3). The pattern and degree of temporal and spatial divergence in the nuclear microsatellites ScAAT1–ScAAT6 among subpopulations was estimated by F_{ST} determination over all loci by

AMOVA analysis, as shown in Table 4. For this analysis, the group of strains obtained from each sampling site in each year was considered as a population. The contribution of variation within the populations defined was always very high, ranging from 81 to 93%, as might be expected from a set of highly polymorphic loci. For the analysis of variation between vineyards and between sampling years, the assemblage of several populations from one vineyard or sampling year was considered as a group. Similarly, for the comparison between sampling sites within a vineyard, each of the sampling sites represented a group of strains that was made up of the populations found in the three sampling years. For all analyses, differences within groups constitute 6.3–24.5%, whereas differences among groups constitute only up to 7% of variation. Populations from C (2002) were not included in this analysis, given that a single genetic pattern was obtained for the spontaneous fermentation of grapes collected from site CIV.

In order to assess whether the occurrence of commercial yeast strains may contribute to the genetic homogeneization of the populations from vineyards A, C and P, calculations were performed including or not genotypes from the recovered commercial yeast strains. Globally,

Table 3 Observed (*Ho*) and expected (*He*) heterozygosity for *S. cerevisiae* populations from vineyards A, C and P

Locus		Vineyard A (94 genotypes)	Vineyard C (70 genotypes)	Vineyard P (140 genotypes)
ScAAT1–29 alleles (12–61 repeats)	*Ho*	0.287	0.186	0.236
	He	0.831	0.839	0.832
ScAAT2–14 alleles (1–16 repeats)	*Ho*	0.191	0.286	0.200
	He	0.836	0.866	0.785
ScAAT3–19 alleles (10–49 repeats)	*Ho*	0.212	0.157	0.286
	He	0.881	0.807	0.840
ScAAT4–17 alleles (6–27 repeats)	*Ho*	0.106	0.114	0.157
	He	0.672	0.619	0.468
ScAAT5–6 alleles (13–30 repeats)	*Ho*	0.170	0.229	0.200
	He	0.713	0.708	0.700
ScAAT6–10 alleles (13–28 repeats)	*Ho*	0.042	0.142	0.136
	He	0.463	0.427	0.393

Table 4 AMOVA analysis, F_{ST} values and distribution of variance components (%) among groups (AG), among populations within groups (APWG), and within populations (WP) based on microsatellite data for defined populations, including or not the genotypes of commercial strains that were found in some of the sampling sites, as indicated in Fig. 1

Source of variation			– commercial strains					+ commercial strains				
			AG	AGWP	WP	F_{ST}	$P(r < o)$	AG	AGWP	WP	F_{ST}	$P(r < o)$
Among vineyards	2001	A/P	3.03	9.03	87.94	0.12	< 0.0001	3.68	6.94	89.39	0.11	< 0.0001
	2002		6.38	13.28	80.33	0.20	0.0001	5.60	11.92	82.48	0.18	< 0.0001
	2003		2.76	11.29	85.95	0.14	0.0001	2.71	10.85	86.44	0.14	< 0.0001
	2001	A/C	– 4.16	16.66	87.51	0.12	0.059	3.91	8.75	87.33	0.13	0.0244
	2003		1.09	16.20	82.71	0.17	< 0.0001	1.55	15.10	83.34	0.17	< 0.0001
	2001	P/C	– 1.21	8.31	92.89	0.07	0.0001	0.64	5.61	93.75	0.06	0.0001
	2003		0.48	8.10	91.42	0.09	< 0.0001	0.03	7.22	92.75	0.07	0.004
Among years	2001/2002	A	– 2.45	13.94	88.51	0.11	0.034	– 2.45	13.94	88.51	0.11	0.03519
		P	0.79	9.94	89.27	0.11	0.0001	– 0.41	7.35	93.06	0.07	0.003
	2002/2003	A	1.29	15.79	83.0	0.17	< 0.0001	1.23	15.55	83.22	0.17	< 0.0001
		P	1.68	7.73	90.59	0.09	0.052	0.01	6.68	93.30	0.07	0.106
	2001/2003	A	– 2.45	20.48	82.05	0.18	< 0.0001	– 2.58	20.01	82.57	0.17	< 0.0001
		C	– 1.56	12.67	88.89	0.11	0.0001	2.20	8.63	89.17	0.11	0.0001
		P	0.37	6.30	93.33	0.07	0.0001	0.15	5.09	94.77	0.05	0.003
Among sampling sites	2001+	A	– 0.02	16.65	83.38	0.17	< 0.0001	0.48	15.99	83.53	0.16	< 0.0001
	2002+	C	– 12.27	24.46	87.81	0.12	0.0001	– 8.31	18.78	89.53	0.10	< 0.0001
	2003	P	– 1.23	9.19	92.05	0.08	< 0.0001	– 0.82	6.88	93.94	0.06	0.0001

and for all analysis performed, F_{ST} values range between 0.05 and 0.20, corresponding to a moderate (0.05–0.15) to great (0.15–0.25) genetic differentiation (Wright 1978). Statistically significant genetic variation (*P*(random value < observed value) < 0.001) was found at every level of analysis (among vineyards, among year-classes). The inclusion of commercial yeast's genotypes found in the three vineyards just slightly reduced the F_{ST} by merely 0.01 to 0.02 values, in about 2/3 of the comparisons performed.

When populations from different vineyards were pair-wise associated (A/C, A/P and P/C), F_{ST} values of the same order of magnitude were found in consecutive years, being higher for A/C and A/P (0.12–0.17 and 0.11–0.20) when compared to P/C (0.06–0.09). Most of the *S. cerevisiae* populations from A, C and P were significantly different in three consecutive years, and populations within a vineyard varied in consecutive years, being more variable in A (F_{ST} = 0.11–0.18) than in P (F_{ST} = 0.05–0.11). When samples were
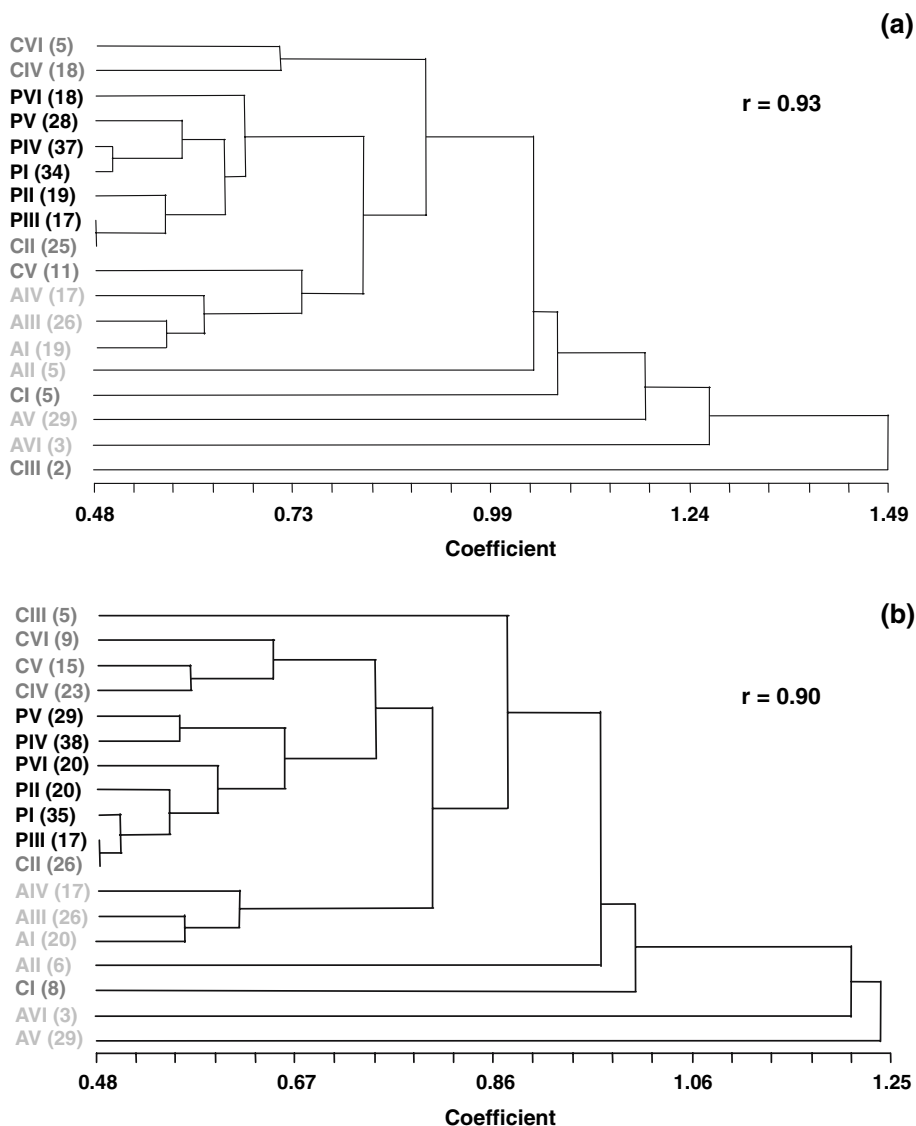
Fig. 3 UPGMA phenogram based on Euclidean distance of allelic frequencies from strains found at each sampling site over 3 years excluding (**a**) or including (**b**) the genotypes of commercial yeast strains. Numbers in parenthesis indicate the number of strains corresponding to unique patterns

pooled across year-classes within the sampling sites of each vinery, the highest F_{ST} value was again obtained for A (0.16–0.17) compared to C (0.10–0.12) and P (0.06–0.08).

Similarity of populations from vineyards A, C and P

Relationships among the populations belonging to six sampling points in three wineries, that were isolated during the 3 years sampling campaigns, were determined by a cluster analysis (UPGMA) based on a Euclidean distance dissimilarity matrix of allelic frequencies (Fig. 3). The cophenetic correlation factor r was 0.93 and 0.90 when genotypes of commercial yeast strains were included or not in this analysis, indicating that the genetic relationships were not distorted by hierarchic clustering. A similar genetic structure was obtained with the neighbor joining algorithm (not shown), being the value for r significantly lower (0.74). For the

analysis performed without commercial yeast's genotypes, populations were grouped in three clusters at a dissimilarity distance of about 0.60–0.65, comprising two sampling sites of C, six sampling sites of P, and three sites of A, showing the existence of a certain populational substructure, characteristic for each vineyard. Population CII lies within the cluster P, and strains isolated from CV are located within the A-cluster, indicating that genetic differences do not delimit specific populations with fixed geographic boundaries.

Further exceptions from a vineyard-specific population structure were found for sampling sites CI, CIII, AII, and AVI, possibly due to the low number of strains and consequent lack of rigor in the quantification of allelic frequencies. Sampling site V in vineyard A is also located outside the A-cluster and showed the most divergent allelic frequencies from all populations, although a sufficient number of strains (27) were analyzed. The high frequency of allele 24 (ScAAT4) in strains collected during 2003 in site V may be the main reason for this observation.

Populations within groups C and P are in general more closely related, and populations from sampling points in vineyard P are more similar to each other as indicated by the dissimilarity distance between them. *S. cerevisiae* populations belonging to vineyard A seem more heterogeneous and also more distinct from C and P. These data are in accordance with the pairwise comparison of vineyards and the respective F_{ST} values as a measure of genetic differentiation, as previously shown in Table 4.

The general structure of the dendrogram was maintained when commercial yeast's genotypes were included. As expected, populations from CIV, CV and CVI are closer related, due to the presence of strains Zymaflore VL1, F10, F15, Uvaferm BDX and Lalvin ICV D254 in these sites located close (10–20 m) to the winery where the strains have been used.

In the present study, genetic distances and geographical localization of the populations did not correlate, since strains with most similar genotypes resided in most distant vineyards C–P (~100 km). The opposite situation was verified for

the closer vineyards A–C (~60 km) and A–P (~40 km) (Fig. 1).

Discussion

Vineyard-associated *S. cerevisiae* populations have never been extensively characterized by microsatellite markers. The initial screening of 1620 isolates by mtDNA RFLP and subsequent microsatellite analysis of 361 strains was revealed to be an appropriate strategy for the present large-scale approach, since both methods are equivalent concerning their capacity to discriminate commercial wine yeast strains (Schuller et al. 2004).

Some remarks have to be made concerning our experimental approach. The isolated *S. cerevisiae* strains may not be truly representative of the vineyard population because strains were isolated after enrichment through must fermentation. Grape must creates selective and very stressful conditions for yeast, totally distinct from the environmental influences in nature and fermentative ability may not be correlated with evolutionary fitness in a vineyard ecosystem. Rarely occurring strains, although able to survive fermentation, might also have not been detected as the detection limit of our experimental approach is 3.3% (one strain in 30 isolates). Using previously proposed direct-plating methods from single grape berries, would be highly labor-intensive and would not permit a search for fermenting yeasts, especially *S. cerevisiae*, in 18 sites, in two campaigns and over 3 years. Therefore we regard our approach as an acceptable compromise, allowing good estimation of population composition, but preventing a precise description in terms of relative strain abundance in nature.

Analysis of microsatellite loci showed a significant excess of homozygotes, the observed heterozygosity was three to four times lower than the estimated value. Heterozygous genotypes reduction relative to that expected under random mating is a consequence of population substructuring. Wine strains of *S. cerevisiae* are usually prototrophic homothallic diploids, mostly homozygous for the homothallism gene (HO/HO) and have high spore viability contrary to strains

with heterozygosities that show decreased spore viabilities with increasing number of heterozygous loci, associated with reduced strain fitness. A mechanism called "genome renewal" (Mortimer et al. 1994) has been proposed for natural wine yeast strains that undergo mating among their progeny cells and thereby change a multiple heterozygote into completely homozygous diploids, leading to gradual replacement of heterozygous diploids. The most likely situation in yeasts is therefore asexual reproduction with some cycles of homothallic self-mating (genome renewal), which would generate the high homozygosity observed. However, an alternative possibility for the high degree of homozygosity observed could be mitotic recombination or gene conversion during asexual reproduction. Heterozygous deficiencies can also be explained by the presence of null alleles that arise when mutations prevent primers from binding, so that many of the apparent homozygotes can be, in reality, heterozygotes between a visible and a null allele. The high degree of homozygosity points to the existence of genetically isolated clonal subpopulations of *S. cerevisiae* strains with distinct genetic constitution. Since a primarily sexual reproduction is not prevalent and the populations are not in equilibrium, further genetic analysis could not be performed.

The dendrogram shown in Fig. 3 and AMOVA analysis (Table 4) clearly agree in the distinction of the more similar populations belonging to vineyard P and C compared to A. Allelic frequencies-based clustering of at least 10 distinct genotypes lead to the expected result concerning populational structures, showing that ecologically meaningful conclusions require an adequate sample size. As most alleles are widespread, certainly due to the relatively close location of the vineyards, genetic differences among *S. cerevisiae* populations derived mainly from gradations in allele frequencies rather than from distinctive "diagnostic" genotypes. Only the accumulation of small allele-frequency differences across six loci allowed the identification of a population structure. Some of the allelic variation may also be linked to loci which determine fermentative ability, which may explain some of the similarities between yeast from different vineyards.

Several commercial yeast strains have been used for the last years in the wineries that are located within the vineyards and were recovered in the present study. The structure of the dendrograms including or not the genotypes of commercial strains is similar, indicating that the closer genetic proximity of populations from C and P is due to autochthonous strains and that the rate of gene flow caused by continuous use of starter yeasts was not sufficient to genetically homogenize local indigenous strains. A detailed analysis about the dynamics and survival of industrial yeast strains in the mentioned vineyards and in three vineyards of the Languedoc wine region in France showed that the asexual dispersal of these strains is very limited (occurring at a distance between 10–200 m from the winery) and is largely favoured by the presence of water runoff. Commercial strains were mostly found in the samples collected after harvest, reflecting their immediate dissemination after wineries started wine production. Permanent implantation in the vineyard did not occur, the strains rather showed natural fluctuations of periodical appearance/disappearance like autochthonous strains (Valero et al. 2005).

In the present study, 52 new alleles were identified besides the 41 alleles previously described for ScAAT1-ScAAT6 (Pérez et al. 2001). In the meantime, other highly polymorphic microsatellite markers have been described for *S. cerevisiae* (Bradbury et al. 2006; Legras et al. 2005). Multiplex amplification of a highly polymorphic set of microsatellites would be desirable and yeast researchers should find common criteria for the generation and storage of microsatellite data of *S. cerevisiae* strains. It is important to indicate alleles as a number of repeats rather than amplicon sizes, because some authors use the same microsatellite markers but distinct primer pairs for their amplification. The extension of the current approach to strains isolated from other viticultural regions is desirable, since a preliminary comparison revealed major differences in both allelic combinations and frequencies (our unpublished data).

The occurrence and survival of *S. cerevisiae* in vineyards depends on numerous factors like climatic influence such as rainfall, temperature (Longo et al. 1991; Parish and Carroll 1985) or

viticultural practices like agrochemical applications, grape variety or maturation stage (Pretorius et al. 1999; Rosini 1982). In the present case, the three geographically close vineyards share climate similarities, but one can not exclude microclimatic influences, not recorded in the present study. Geographical distance was not correlated with genetic proximity, since the most distant (100 km) vineyards P and C had most similar populations. This is coincident with data of previous studies (Torija et al. 2001; Versavaud et al. 1995), but it was also shown that this correlation exists among *S. cerevisiae* strains from different Spanish wine regions, being red wine strains significantly grouped according to their geographic origin, independently of the wine type and the grapevine cultivar, and white wine strains according to ecological factors such as wine type of grapevine cultivars (Guillamon et al. 1996). The three sampled sub-regions share similar viticultural practices, being Loureiro the grape variety of vineyard A, Alvarinho and Avesso the cultivars of vineyard P and C, respectively. Correlation between grape variety and global genetic constitution of associated strains seems tempting, but more experimental data are needed to support such a hypothesis.

Genetic differentiation (the acquisition of allele frequencies that differ among subpopulations) may result from natural selection favoring different genotypes in different subpopulations, but it may also result from random processes in the transmission of alleles from one generation to the next or from stochastic differences in allele frequency among the initial founders of the subpopulations. The distinction between little ($F_{ST} = 0$–0.05), moderate ($F_{ST} = 0.05$–0.15), great ($F_{ST} = 0.15$–0.25) and very great ($F_{ST} > 0.25$) genetic differentiation has been suggested (Wright 1978), but the identification of causes underlying a particular F_{ST} value can be difficult. AMOVA analysis revealed to be useful for the detection of inter-populational genetic variations among populations that exhibit a high amount of intra-populational variability. Genetic differentiation among populations grouped according to sampling year or site, being the highest value recorded for vineyard A, followed by C and P. Differences in the same vineyard in consecutive years are of the same order of magnitude as the differences verified among the three vineyards, demonstrating the importance of sampling in consecutive years in order to get a realistic picture of yeast population distribution. Differences over time that are the same as differences over distance could result from slightly detrimental alleles (or mutations) that are being selectively removed from the population or from a population going through a series of bottlenecks (e.g. the time from the end of one season to the beginning of the next) that result in differences in gene frequencies due to drift. Values of genetic differentiation are correlated with the distance between sampling points and consequently the size of the vineyards. *S. cerevisiae* strains may become more distinctive in a larger vineyard that constitutes a bigger "evolutionary playground", hypothesizing that local populations may evolve due to multi-factorial influences being the size of the vineyard one of them. Genetic heterogeneity in a vine could follow a pattern of isolation-by-distance, where genetic divergence increases with vineyard size. However, the forces causing a global shift in a vineyard's *S.cerevisiae* population still remain to be clarified.

The present work is to our knowledge the first large-scale approach addressing the usefulness of microsatellite typing in an ecological survey of indigenous *S. cerevisiae* strains isolated from vineyards. Microsatellite typing with loci ScAAT1–ScAAT6, followed by statistical analysis permitted a very fine population screen, and is therefore the appropriate method to obtain deeper insight in ecology and biogeography of *S. cerevisiae* strains, even among geographically close regions. These studies are indispensable for developing strategies aiming at the preservation of biodiversity and genetic resources as a basis for further strain selection.

References

Beltran G, Torija MJ, Novo M, Ferrer N, Poblet M, Guillamon JM, Rozes N, Mas A (2002) Analysis of yeast populations during alcoholic fermentation: a six year follow-up study. Syst Appl Microbiol 25:287–293

Blondin B, Vezinhet F (1988) Identification de souches de levures oenologiques par leurs caryotypes obtenus en électrophorèse en champs pulsée. Rev Fr Oenol 28:7–11

Bradbury J, Richards K, Niederer H, Lee S, Rod Dunbar P, Gardner R (2006) A homozygous diploid subset of commercial wine yeast strains. Antonie van Leeuwenhoek 89:27–38

Briones AI, Ubeda JF, Cabezudo MD, Martin-Alvarez P (1995) Selection of spontaneous strains of Saccharomyces cerevisiae as starters in their viticultural area. In: Charalambous G (ed) Food flavours: generation, analysis and process influence. Elsevier Science, Amsterdam, pp 1597–1622

Constanti M, Poblet M, Arola L, Mas A, Guillamon JM (1997) Analysis of yeast populations during alcoholic fermentation in a newly established winery. Am J Enol Vitic 48:339–344

Farris GA, Budroni M, Vodret T, Deiana P (1990) Sull'originr dei lieviti vinari i lieviti dei terreni, delle foglie e degli acini di alcun vigneti sardi. L'Enotecnico 6:99–108

Fleet GH, Heard GM (1993) Yeasts: growth during fermentation. In: Fleet GH (ed) Wine microbiology and biotechnology. Harwood, pp 27–55

Frezier V, Dubourdieu D (1992) Ecology of yeast strain Saccharomyces cerevisiae during spontaneous fermentation in a Bordeaux winery. Am J Enol Vitic 43:375–380

Gallego FJ, Perez MA, Martinez I, Hidalgo P (1998) Microsatellites obtained from database sequences are useful to characterize Saccharomyces cerevisiae strains. Am J Enol Vitic 49:350–351

Guillamon JM, Barrio E, Querol A (1996) Characterization of wine yeast strains of the Saccharomyces genus on the basis of molecular markers: relationships between genetic distance and geographic or ecological origin. Syst Appl Microbiol 19:122–132

Hennequin C, Thierry A, Richard GF, Lecointre G, Nguyen HV, Gaillardin C, Dujon B (2001) Microsatellite typing as a new tool for identification of Saccharomyces cerevisiae strains. J Clin Microbiol 39:551–559

Legras JL, Karst F (2003) Optimisation of interdelta analysis for Saccharomyces cerevisiae strain characterisation. FEMS Microbiol Lett 221:249–255

Legras JL, Ruh O, Merdinoglu D, Karst F (2005) Selection of hypervariable microsatellite loci for the characterization of Saccharomyces cerevisiae strains. Int J Food Microbiol 102:73–83

Longo E, Cansado J, Agrelo D, Villa TG (1991) Effect of climatic conditions on yeast diversity in grape musts from northwest Spain. Am J Enol Vitic 42:141–144

Lopez V, Querol A, Ramon D, Fernandez-Espinar MT (2001) A simplified procedure to analyse mitochondrial DNA from industrial yeasts. Int J Food Microbiol 68:75–81

Martini A, Ciani M, Scorzetti G (1996) Direct enumeration and isolation of wine yeasts from grape surfaces. Am J Enol Vitic 47:435–440

Martini A, Frederichi F, Rosini G (1980) A new approach to the study of yeast ecology on natural substrates. Can J Microbiol 26:856–859

Monteil H, Blazy-Mangen F, Michel G (1986) Influence des pesticides sur la croissance des levures des raisins et des vins. Sci Aliments 6:349–360

Mortimer R, Polsinelli M (1999) On the origins of wine yeast. Res Microbiol 150:199–204

Mortimer RK, Romano P, Suzzi G, Polsinelli M (1994) Genome renewal—a new phenomenon revealed from a genetic study of 43 strains of Saccharomyces cerevisiae derived from natural fermentation of grape musts. Yeast 10:1543–1552

Naumov GI, James SA, Naumova ES, Louis EJ, Roberts IN (2000) Three new species in the Saccharomyces sensu stricto complex: Saccharomyces cariocanus, Saccharomyces kudriavzevii and Saccharomyces mikatae. Int J Syst Evol Microbiol 50:1931–1942

Ness F, Lavallée F, Dubourdieu D, Aigle M, Dulau L (1993) Identification of yeast strains using the polymerase chain reaction. J Sci Food Agric 62:89–94

Parish ME, Carroll DE (1985) Indigenous yeasts associated with muscadine (Vitis rotundifolia) grapes and musts. Am J Enol Vitic 36:165–169

Pérez MA, Gallego FJ, Martinez I, Hidalgo P (2001) Detection, distribution and selection of microsatellites (SSRs) in the genome of the yeast Saccharomyces cerevisiae as molecular markers. Lett Appl Microbiol 33:461–466

Pretorius IS, van der Westhuizen TJ, Augustyn OHP (1999) Yeast biodiversity in vineyards and wineries and its importance to the South African wine industry. S Afr J Enol Vitic 20:61–74

Querol A, Barrio E, Huerta T, Ramon D (1992) Molecular monitoring of wine fermentations conducted by active dry yeast strains. Appl Environ Microbiol 58:2948–2953

Redzepovic S, Orlic S, Sikora S, Majdak A, Pretorius IS (2002) Identification and characterization of Saccharomyces cerevisiae and Saccharomyces paradoxus strains isolated from Croatian vineyards. Lett Appl Microbiol 35:305–310

Regodon JA, Perez F, Valdes ME, DeMiguel C, Ramirez M (1997) A simple and effective procedure for selection of wine yeast strains. Food Microbiol 14:247–254

Rosini G (1982) Influenza della microflora saccaromicetico della cantina sulla fermentazione del mosto d'uva. Vigne Vini 9:43–46

Sabate J, Cano J, Esteve-Zarzoso B, Guillamon JM (2002) Isolation and identification of yeasts associated with vineyard and winery by RFLP analysis of ribosomal

genes and mitochondrial DNA. Microbiol Res 157:267–274

Schneider S, Roessli D, Excoffier L (1997) Arlequin ver 2.000: a software for population genetic data analysis. Genetics and Biometry Laboratory, Department of Anthropology and Ecology, University of Geneva, Switzerland

Schuller D, Alves H, Dequin S, Casal M (2005) Ecological survey of Saccharomyces cerevisiae strains from vineyards in the Vinho Verde Region of Portugal. FEMS Microbiol Ecol 51:167–177

Schuller D, Valero E, Dequin S, Casal M (2004) Survey of molecular methods for the typing of wine yeast strains. FEMS Microbiol Lett 231:19–26

Torija MJ, Rozes N, Poblet M, Guillamon JM, Mas A (2001) Yeast population dynamics in spontaneous fermentations: Comparison between two different wine-producing areas over a period of three years. Antonie van Leeuwenhoek 79:345–352

Valero E, Schuller D, Gambon B, Casal M, Dequin S (2005) Dissemination and survival of commercial wine yeast in the vineyard: a large-scale, three-years study. FEMS Yeast Res 5:959–969

Vaughan-Martini A, Martini A (1995) Facts, myths and legends on the prime industrial microorganism. J Ind Microbiol 14:514–522

Versavaud A, Courcoux P, Roulland C, Dulau L, Hallet J-N (1995) Genetic diversity and geographical distribution of wild Saccharomyces cerevisiae strains from the wine-producing area of Charentes, France. Appl Environ Microbiol 61:3521–3529

Wright S (1978) Evolution and the genetics of populations, vol. 4. University of Chicago Press, Chicago