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

Dorit Schuller (*) and Margarida Casal

Centro de Biologia (CB-UM), Departamento de Biologia, Universidade do Minho, 4710-057 Braga, Portugal

Keywords: microsatellite; *S. cerevisiae*; commercial yeasts; yeast population

*Corresponding author:

Dorit Schuller

Centro de Biologia, Departamento de Biologia, Universidade do Minho

Campus de Gualtar

4710-057 Braga, Portugal

Phone: +351 253 604310

Fax: +351 253 678980

email: dschuller@bio.uminho.pt

Abstract

From the analysis of six polymorphic microsatellite loci performed in 361 *Saccharomyces cerevisiae* isolates 93 alleles were identified, being 52 of them 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 three to four 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 diferent 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.

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 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 nowadays commercialized as active dry yeast. Such strains are capable 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 nowadays 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 6 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 (10-200m) 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 to find 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., 2005; 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).

Aiming at 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 *Saccharomyces 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 Figure 1. In each vineyard, six sampling points were defined, located at ten to 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, whereas 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 20ºC 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, 28ºC). The isolates obtained throughout this work were stored in glycerol (30%, v/v) at -80ºC.

135 DNA isolation

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

141

142 Mitochondrial DNA restriction patterns

143 Mitochondrial DNA restriction of all strains was carried out as a first screening
144 approach, to reduce the number of isolates to be analysed by microsatellite typing.
145 Digestion reactions were carried out overnight at 37ºC and contained 15 µl of the
146 previously isolated DNA, and were prepared as previously described (Schuller et al.,
147 2004), in a final volume of 20 µl. To each isolate a pattern designation was attributed
148 (A1-A92, C1-C70 and P1-P135 for isolates from vineyard A, C and P respectively).
149 When isolates from different samples showed identical patterns, one representative
150 strain from each sample was randomly withdrawn, resulting in a total of 361 isolates
151 that were further studied by microsatellite analysis.

152

153 Microsatellite amplification

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

157

158 Computer assisted analysis

159 Based on the the genome sequence for strain S288C (SGD database, http://genome-
160 www.stanford.edu.saccharomyces), and the results obtained for the size of
161 microsatellite amplicons of this strain, the number of repeats for alleles from each locus
162 was calculated. Genetic analysis was performed using the software Arlequin 2000
163 (Schneider et al., 1997) and included (i) estimation of allelic frequencies (ii) observed
164 heterozygosity compared to expected values, (iii) estimation of Wright's F_{ST} value
165 (Wright, 1978) and (iv) genetic variation attributable to different hierarchical levels of
166 defined genetic structures (AMOVA analysis). Wright's F_{ST} value was calculated to
167 determine population differentiation among vineyards, among sampling years and also
168 among sampling locations within a vineyard.
169 An allelic frequencies matrix was obtained based on Euclidean distance and clustered
170 by the unweighted pair group method arithmetic mean (UPGMA) using the program
171 NTSYSpc 2.0 (Applied Biostatistics Inc.) to examine whether genetic divergence was
172 correlated with sampling sites. This software was also used for dendrogram drawing and
173 to calculate a cophenetic correlation coefficient (r).
174
175 **Results**
176 Obtention of *S. cerevisiae* strains
177 As shown in Figure 1, six sampling sites in each of three vineyards, located in the
178 Vinho Verde Wine Region, were sampled during the 2001-2003 harvest seasons. Two
179 sampling campaigns were performed, one before and another after the harvest, in a time
180 frame of about two weeks as an attempt to obtain an elevated number of different
181 strains. A total of 108 grape samples have been planned (six sampling points x two
182 sampling campaigns x three vineyards x three years), from which 54 started a

183 spontaneous fermentation, 36 were not able to start fermentation after 30 days of
184 incubation, whereas 18 samples were not collected due to unfavorable weather
185 conditions and a bad sanitation state of the grapes in 2002. From the 54 fermentations
186 1620 yeast isolates were obtained. All the isolates were analyzed by their mtDNA RFLP
187 (*Hinf*I) and a pattern profile was attributed to each isolate, resulting in a total of 297
188 different profiles. The results of this ecological survey, including the temporal and
189 spatial distribution of the found strains has been recently published (Schuller et al.,
190 2005). When the same profile was found in more than one sample, one strain from each
191 sample was randomly withdrawn resulting in a total of 361 isolates, all assumed to be *S.*
192 *cerevisiae* strains. This was supported by their inability to grow in a medium containing
193 lysine as sole nitrogen source and by their capacity to amplify the previously described
194 *S. cerevisiae* specific microsatellite loci ScAAT1 – ScAAT6 (Pérez et al., 2001).
195 The species *S. cerevisiae* is very closely related to the species *Saccharomyces bayanus*,
196 *Saccharomyces pastorianus*, *Saccharomyces paradoxus*, *Saccharomyces cariocanus*,
197 *Saccharomyces mikatae*, and *Saccharomyces kudriavzevii* (Naumov et al., 2000). These
198 six species, together with *S. cerevisiae*, constitute the *Saccharomyces* sensu stricto
199 complex. Only *S. cerevisiae*, *S. bayanus*, *S. pastorianus*, and *S. paradoxus* are
200 associated with fermentative processes. *S. cerevisiae* and *S. bayanus* are considered the
201 predominating species in wine fermentation. *S. paradoxus* has been isolated only once
202 in wine (Redzepovic et al., 2002), whereas *S. pastorianus* is only present in beer
203 making. Our (unpublished) results showed that the specific microsatellite primers are
204 not amplifying the homologous loci from other *Saccharomyces* species such as *S.*
205 *bayanus* and *S. paradoxus.* Sequence analysis was performed with data obtained from
206 the Washington University Genome Sequencing Center

207 (http://genome.wustl.edu/projects/yeast/) and the Broad Institute
208 (http://www.broad.mit.edu/annotation/fungi/comp_yeasts/). Both *S. bayanus* and *S.*
209 *paradoxus* showed no homology with the ScAAT1, ScAAT3, and ScAAT6 primer
210 binding regions. ScAAT4, ScAAT5 and ScAAT2, ScAAT5 primer binding sites had a
211 low homology with the corresponding sequences in *S. bayanus* and *S. paradoxus*,
212 respectively.
213 Strains showing different mtDNA RFLP patterns had distinct genotypes as determined
214 by the allelic combinations for loci ScAAT1-ScAAT6. Microsatellite analysis
215 performed in a ramdomly seleccted group of 50 isolates (among the whole collection
216 comprising 1620 strains) showed that isolates with the same/different microsatellite
217 amplification profiles always corresponded to the same/different mtDNA RFLP
218 patterns. In addition, 90 isolates with identical mtDNA RFLP were analyzed in 6
219 microsatellite loci and always showed the same allelic combinations (our unpublished
220 results). Therefore, allele frequencies correspond to a random sampling of the alleles
221 present in the microfermentations.
222 The table in Figure 1 indicates the number of different microsatellite genotypes obtained
223 from strains collected at each sampling site in both sampling campaigns (before and
224 after the harvest). The number of different strains isolated from each sampling point
225 showed a lower (one to ten strains) or higher (11 - 21 strains) biodiversity. Genotypes a-
226 k showed a wider temporal and geographical distribution, being the corresponding
227 strains characterized by a generalized pattern of sporadic presence, absence and
228 reappearance across sampling sites, vineyards or years. Genotype b showed a more
229 regional distribution with a perennial behavior. In several sampling sites commercial
230 strains were recovered, that have been used predominately (in higher quantity and

231 continuously) or sporadically (in lower quantity and not continuously) by the wineries
232 during the harvests preceeding the 5 years of the current study. The respective
233 genotypes are shown in Table 2. A detailed analysis regarding their predominance and
234 spatio-temporal distribution, including also the results from an identical study
235 performed in the Languedoc wine region (France) has been recently published (Valero
236 et al., 2005).

237

238 Genetic analysis of alleles obtained for loci ScAAT1 – ScAAT6

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

250 Populations from C and P share the most frequent alleles for markers ScAAT1,
251 ScAAT2 and ScAAT3 (17, 14 and 22), while populations belonging to A had the
252 highest frequencies at alleles 28, 13 and 20, respectively. For ScAAT4 and ScAAT6,
253 alleles 20 and 16 were the most frequent for all 3 populations, and for locus ScAAT5
254 the allele 16 was most frequent in A and C, and allele 15 in P respectively. Many of the

255 alleles occurring with a lower global frequency, showed different incidences for *S.*
256 *cerevisiae* populations from vineyards A, C and P (e.g. allele 26 and 27, ScAAT1; allele
257 7, 11 and 12, ScAAT2; allele 17 and 23, ScAAT3; allele 24, ScAAT4; allele 17,
258 ScAAT5; allele 17; ScAAT6). For each locus, unique alleles were also found in each of
259 the three populations; their frequencies were very low, ranging between 0.01 and 0.03,
260 and they might play only a minor role.
261 For the populations from different vineyards the observed heterozygosity (*Ho*) was in
262 general about three to four times lower than the expected heterozygosity (*He*) for all loci
263 analyzed (Table 3). The pattern and degree of temporal and spatial divergence in the
264 nuclear microsatellites ScAAT1 to ScAAT6 among subpopulations was estimated by
265 F_{ST} determination over all loci by AMOVA analysis, as shown in Table 4. For this
266 analysis, the group of strains obtained from each sampling site in each year was
267 considered as a population. The contribution of variation within the populations defined
268 was always very high, ranging from 81 to 93%, as might be expected from a set of
269 highly polymorphic loci. For the analysis of variation between vineyards and between
270 sampling years, the assemblage of several populations from one vineyard or sampling
271 year was considered as a group. Similarly, for the comparison between sampling sites
272 within a vineyard, each of the sampling sites represented a group of strains that was
273 made up of the populations found in the 3 sampling years. For all analysis, differences
274 within groups constitute 6.3 to 24.5%, whereas differences among groups constitute
275 only up to 7% of variation. Populations from C (2002) are not included in this analysis,
276 given that a single genetic pattern was obtained for the spontaneous fermentation of
277 grapes collected from site CIV.

278 In order to assess whether the occurrence of commercial yeast strains may contribute to
279 the genetic homogeneization of the populations from vineyards A, C and P, calculations
280 were performed including or not genotypes from the recovered commercial yeast
281 strains. Globally, and for all analysis performed, F_{ST} values range between 0.05 and
282 0.20, corresponding to a moderate (0.05 – 0.15) to great (0.15 – 0.25) genetic
283 differentiation (Wright, 1978). Statistically significant genetic variation (P(random
284 value< observed value) < 0.001) was found at every level of analysis (among vineyards,
285 among year-classes). The inclusion of commercial yeast's genotypes found in the 3
286 vineyards just slightly reduced the F_{ST} by merely 0.01 to 0.02 values, in about 2/3 of the
287 comparisons performed.
288 When populations from different vineyards were pair wise associated (A/C, A/P and
289 P/C), F_{ST} values of the same order of magnitude were found in consecutive years, being
290 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).
291 Most of the *S. cerevisiae* populations from A, C and P were significantly different in
292 three consecutive years, and populations within a vineyard varied in consecutive years,
293 being more variable in A (F_{ST} = 0.11 – 0.18) than in P (F_{ST} = 0.05 – 0.11). When
294 samples were pooled across year-classes within the sampling sites of each vinery, the
295 highest F_{ST} value was again obtained for A (0.16 - 0.17) compared to C (0.10 – 0.12)
296 and P (0.06 - 0.08).
297
298 Similarity of populations from vineyards A, C and P
299 Relationships among the populations belonging to six sampling points in three wineries,
300 that were isolated during the 3 years sampling campaigns, were determined by a cluster
301 analysis (UPGMA) based on a Euclidean distance dissimilarity matrix of allelic

302 frequencies (Figure 3). The cophenetic correlation factor r was 0.93 and 0.90 when
303 genotypes of commercial yeast strains were included or not in this analysis, indicating
304 that the genetic relationships were not distorted by hierarchic clustering. A similar
305 genetic structure was obtained with the neighbor joining algorithm (not shown), being
306 the value for r significantly lower (0.74). For the analysis performed without
307 commercial yeast's genotypes, populations were grouped in three clusters at a
308 dissimilarity distance of about 0.60 – 0.65, comprising two sampling sites of C, six
309 sampling sites of P, and three sites of A, showing the existence of a certain populational
310 substructure, characteristic for each vineyard. Population CII lies within the cluster P,
311 and strains isolated from CV are located within the A-cluster, indicating that genetic
312 differences do not delimit specific populations with fixed geographic boundaries.
313 Further exceptions from a vineyard - specific population structure were found for
314 sampling sites CI, CIII, AII, and AVI, possibly due to the low number of strains and
315 consequent lack of rigor in the quantification of allelic frequencies. Sampling site V in
316 vineyard A is also located outside the A-cluster and showed the most divergent allelic
317 frequencies from all populations, although a sufficient number of strains (27) were
318 analyzed. The high frequency of allele 24 (ScAAT4) in strains collected during 2003 in
319 site V may be the main reason for this observation.
320 Populations within groups C and P are in general more closely related, and populations
321 from sampling points in vineyard P are more similar to each other as indicated by the
322 dissimilarity distance between them. *S. cerevisiae* populations belonging to vineyard A
323 seem more heterogeneous and also more distinct from C and P. These data are in
324 accordance with the pairwise comparison of vineyards and the respective F_{ST} values as a
325 measure of genetic differentiation, as previously shown in Table 4.

326 The general structure of the dendrogram was maintained when commercial yeast's
327 genotypes were included. As expected, populations from CIV, CV and CVI are closer
328 related, due to the presence of strains Zymaflore VL1, F10, F15, Uvaferm BDX and
329 Lalvin ICV D254 in these sites located close (10-20 m) to the winery where the strains
330 have been used.
331 In the present study, genetic distances and geographical localization of the populations
332 did not correlate, since strains with most similar genotypes resided in most distant
333 vineyards C - P (~ 100 km). The opposite situation was verified for the closer vineyards
334 A - C (~ 60 km) and A - P (~ 40 km) (Figure 1).
335
336 **Discussion**
337 Vineyard–associated *S. cerevisiae* populations have never been extensively
338 characterized by microsatellite markers. The initial screening of 1620 isolates by
339 mtDNA RFLP and subsequent microsatellite analysis of 361 strains revealed to be an
340 appropriate strategy for the present large-scale approach, since both methods are
341 equivalent concerning their capacity to discriminate commercial wine yeast strains
342 (Schuller et al., 2004).
343 Some remarks have to be made concerning our experimental approach. The isolated *S.*
344 *cerevisiae* strains may not be truly representative of the vineyard population because
345 strains were isolated after enrichment through must fermentation. Grape must creates
346 selective and very stressful conditions for yeast, totally distinct from the environmental
347 influences in nature and fermentative ability may not be correlated with evolutionary
348 fitness in a vineyard ecosystem. Rarely occurring strains, although capable to survive
349 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 to search for fermenting yeasts, especially *S. cerevisiae*, in 18 sites, in two campaigns and over three 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 prevailing and the populations are not in equilibrium, further genetic analysis could not be performed.

The dendrogram shown in Figure 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

398 the asexual dispersal of these strains is very limited (occurring at a distance between 10-
399 200 m from the winery) and is largely favoured by the presence of water runoff.
400 Commercial strains were mostly found in the samples collected after harvest, reflecting
401 their immediate dissemination after wineries started wine production. Permanent
402 implantation in the vineyard did not occur, the strains rather showed natural fluctuations
403 of periodical appearance/disappearance like autochthonous strains (Valero et al., 2005).
404 In the present study, 52 new alleles were identified besides the 41 alleles previously
405 described for ScAAT1-ScAAT6 (Pérez et al., 2001). In the meantime, other highly
406 polymorphic microsatellite markers have been described for *S. cerevisiae* (Bradbury et
407 al., 2005; Legras et al., 2005). Multiplex amplification of a highly polymorphic set of
408 microsatellites would be desirable and yeast researchers should find common criteria for
409 the generation and storage of microsatellite data of *S. cerevisiae* strains. It is important
410 to indicate alleles as a number of repeats rather than amplicon sizes, because some
411 authors use the same microsatellite markers but distinct primer pairs for their
412 amplification. The extension of the current approach to strains isolated from other
413 viticultural regions is desirable, since a preliminary comparison revealed major
414 differences in both allelic combinations and frequencies (our unpublished data).
415 The occurrence and survival of *S. cerevisiae* in vineyards depends on numerous factors
416 like climatic influence such as rainfall, temperature (Longo et al., 1991; Parish and
417 Carroll, 1985) or viticultural practices like agrochemical applications, grape variety or
418 maturation stage (Pretorius et al., 1999; Rosini, 1982). In the present case, the three
419 geographically close vineyards share climate similarities, but one can not exclude
420 microclimatic influences, not recorded in the present study. Geographical distance was
421 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 3 vineyards, demonstrating the importance of sampling in

446 consecutive years in order to get a realistic picture of yeast population distribution.
447 Differences over time that are the same as differences over distance could result from
448 slightly detrimental alleles (or mutations) that are being selectively removed from the
449 population or from a population going through a series of bottlenecks (e.g. the time
450 from the end of one season to the beginning of the next) that results in differences in
451 gene frequencies due to drift. Values of genetic differentiation are correlated with the
452 distance between sampling points and consequently the size of the vineyards. *S.*
453 *cerevisiae* strains may become more distinctive in a larger vineyard that constitutes a
454 bigger "evolutionary playground", hypothesizing that local populations may evolve due
455 to multi-factorial influences being the size of the vineyard one of them. Genetic
456 heterogeneity in a vine could follow a pattern of isolation-by-distance, where genetic
457 divergence increases with vineyard size. However, the forces causing a global shift in a
458 vineyard's *S.cerevisiae* population still remain to be clarified.
459 The present work is to our knowledge the first large-scale approach about the usefulness
460 of microsatellite typing in an ecological survey of indigenous *S. cerevisiae* strains
461 isolated from vineyards. Microsatellite typing with loci ScAAT1-ScAAT6, followed by
462 statistical analysis permitted a very fine population screen, and is therefore the
463 appropriate method to obtain deeper insight in ecology and biogeography of *S.*
464 *cerevisiae* strains, even among geographically close regions. These studies are
465 indispensable for developing strategies aiming at the preservation of biodiversity and
466 genetic resources as a basis for further strain selection.
467
468 **Acknowledgements**

469 This study was supported by the project ENOSAFE (Nº 762, Programa AGRO, medida
470 8), the programme POCI 2010 (project POCI/AGR/56771/2004) and the grant nº 657
471 C2 from the cooperation agreement between the Portuguese Institute for International
472 Scientific and Technological Cooperation (ICCTI) and the French Embassy in Lisbon.
473 We wish to thank Prof. Célia Pais and for helpful discussions during the preparation of
474 this manuscript. Prof. Paula Sampaio is kindly thanked for assistance with statistical
475 analysis. Magda Silva Graça is gratefully acknowledged for the operation of the DNA
476 sequencer. We also appreciate the kind assistance of the enologists Rui Cunha, Anselmo
477 Mendes, Euclides Rodrigues and José Domingues for facilitating sampling campaigns
478 in the three vineyards.

479

480 References

481

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

485 Blondin B, and Vezinhet, F. 1988. Identification de souches de levures oenologiques
486 par leurs caryotypes obtenus en électrophorèse en champs pulsée. Rev. Fr. Oenol.
487 28: 7-11.

488 Bradbury J, Richards, K, Niederer, H, Lee, S, Rod Dunbar, P, and Gardner, R. 2005. A
489 homozygous diploid subset of commercial wine yeast strains. Antonie van
490 Leeuwenhoek (in press) DOI: 10.1007/s10482-005-9006-1.

491 Briones AI, Ubeda, JF, Cabezudo, MD, and Martin-Alvarez, P. 1995. Selection of
492 spontaneous strains of *Saccharomyces cerevisiae* as starters in their viticultural
493 area. In: G. Charalambous (ed.), Food Flavours: Generation, Analysis and Process
494 Influence. Elsevier Science, Amsterdam, pp. 1597-1622.

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

498 Farris GA, Budroni, M, Vodret, T, and Deiana, P. 1990. Sull'originr dei lieviti vinari i
499 lieviti dei terreni, delle foglie e degli acini di alcun vigneti sardi. L'Enotecnico 6:
500 99-108.

501 Fleet GH, and Heard, GM. 1993. Yeasts: growth during fermentation. In: G.H. Fleet
502 (ed.), Wine Microbiology and Biotechnology, Harwood, pp. 27-55.

503 Frezier V, and Dubourdieu, D. 1992. Ecology of yeast strain *Saccharomyces cerevisiae*
504 during spontaneous fermentation in a Bordeaux winery. Am. J. Enol. Vitic. 43:
505 375-380.
506 Gallego FJ, Perez, MA, Martinez, I, and Hidalgo, P. 1998. Microsatellites obtained
507 from database sequences are useful to characterize *Saccharomyces cerevisiae*
508 strains. Am. J. Enol. Vitic. 49: 350-351.
509 Guillamon JM, Barrio, E, and Querol, A. 1996. Characterization of wine yeast strains of
510 the *Saccharomyces* genus on the basis of molecular markers: Relationships
511 between genetic distance and geographic or ecological origin. Syst. Appl.
512 Microbiol. 19: 122-132.
513 Hennequin C, Thierry, A, Richard, GF, Lecointre, G, Nguyen, HV, Gaillardin, C, and
514 Dujon, B. 2001. Microsatellite typing as a new tool for identification of
515 *Saccharomyces cerevisiae* strains. J. Clin. Microbiol. 39: 551-559.
516 Legras JL, and Karst, F. 2003. Optimisation of interdelta analysis for *Saccharomyces*
517 *cerevisiae* strain characterisation. FEMS Microbiol. Lett. 221: 249-255.
518 Legras JL, Ruh, O, Merdinoglu, D, and Karst, F. 2005. Selection of hypervariable
519 microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains Int.
520 J. Food Microbiol. 102: 73-83.
521 Longo E, Cansado, J, Agrelo, D, and Villa, TG. 1991. Effect of climatic conditions on
522 yeast diversity in grape musts from northwest Spain. Am. J. Enol. Vitic. 42: 141-
523 144.
524 Lopez V, Querol, A, Ramon, D, and Fernandez-Espinar, MT. 2001. A simplified
525 procedure to analyse mitochondrial DNA from industrial yeasts. Int. J. Food
526 Microbiol. 68: 75-81.

527 Martini A, Ciani, M, and Scorzetti, G. 1996. Direct enumeration and isolation of wine
528 yeasts from grape surfaces. Am. J. Enol. Vitic. 47: 435-440.
529 Martini A, Frederichi, F, and Rosini, G. 1980. A new approach to the study of yeast
530 ecology on natural substrates. Can. J. Microbiol. 26: 856-859.
531 Monteil H, Blazy-Mangen, F, and Michel, G. 1986. Influence des pesticides sur la
532 croissance des levures des raisins et des vins. Science des Aliments 6: 349-360.
533 Mortimer R, and Polsinelli, M. 1999. On the origins of wine yeast. Res. Microbiol. 150:
534 199-204.
535 Mortimer RK, Romano, P, Suzzi, G, and Polsinelli, M. 1994. Genome renewal - a new
536 phenomenon revealed from a genetic study of 43 strains of *Saccharomyces*
537 *cerevisiae* derived from natural fermentation of grape musts. Yeast 10: 1543-
538 1552.
539 Naumov GI, James, SA, Naumova, ES, Louis, EJ, and Roberts, IN. 2000. Three new
540 species in the Saccharomyces sensu stricto complex: Saccharomyces cariocanus,
541 Saccharomyces kudriavzevii and Saccharomyces mikatae. International Journal of
542 Systematic and Evolutionary Microbiology 50: 1931-1942.
543 Ness F, Lavallée, F, Dubourdieu, D, Aigle, M, and Dulau, L. 1993. Identification of
544 yeast strains using the polymerase chain reaction. J. Sci. Food Agric. 62: 89-94.
545 Parish ME, and Carroll, DE. 1985. Indigenous yeasts associated with muscadine (*Vitis*
546 *rotundifolia*) grapes and musts. Am. J. Enol. Vitic. 36: 165-169.
547 Pérez MA, Gallego, FJ, and Hidalgo, P. 2001. Evaluation of molecular techniques for
548 the genetic characterization of *Saccharomyces cerevisiae* strains. FEMS
549 Microbiol. Lett. 205: 375-378.

550 Pérez MA, Gallego, FJ, Martinez, I, and Hidalgo, P. 2001. Detection, distribution and
551 selection of microsatellites (SSRs) in the genome of the yeast *Saccharomyces*
552 *cerevisiae* as molecular markers. Lett. Appl. Microbiol. 33: 461-466.
553 Pretorius IS, van der Westhuizen, TJ, and Augustyn, OHP. 1999. Yeast biodiversity in
554 vineyards and wineries and its importance to the South African wine industry. S.
555 Afr. J. Enol. Vitic. 20: 61-74.
556 Querol A, Barrio, E, Huerta, T, and Ramon, D. 1992. Molecular monitoring of wine
557 fermentations conducted by active dry yeast strains. Appl. Env. Microbiol. 58:
558 2948-2953.
559 Redzepovic S, Orlic, S, Sikora, S, Majdak, A, and Pretorius, IS. 2002. Identification and
560 characterization of Saccharomyces cerevisiae and Saccharomyces paradoxus
561 strains isolated from Croatian vineyards. Lett. Appl. Microbiol. 35: 305-310.
562 Regodon JA, Perez, F, Valdes, ME, DeMiguel, C, and Ramirez, M. 1997. A simple and
563 effective procedure for selection of wine yeast strains. Food Microbiology 14:
564 247-254.
565 Rosini G. 1982. Influenza della microflora saccaromicetico della cantina sulla
566 fermentazione del mosto d'uva. Vigne Vini 9: 43-46.
567 Sabate J, Cano, J, Esteve-Zarzoso, B, and Guillamon, JM. 2002. Isolation and
568 identification of yeasts associated with vineyard and winery by RFLP analysis of
569 ribosomal genes and mitochondrial DNA. Microbiol. Res. 157: 267-274.
570 Schneider S, Roessli, D, and Excoffier, L. 1997. Arlequin ver 2.000: A software for
571 population genetic data analysis. Genetics and Biometry Laboratory, Department
572 of Anthropology and Ecology, University of Geneva, Switzerland.

573 Schuller D, Alves, H, Dequin, S, and Casal, M. 2005. Ecological survey of
574 *Saccharomyces cerevisiae* strains from vineyards in the Vinho Verde Region of
575 Portugal. FEMS Microbiol. Ecol. 51: 167-177.
576 Schuller D, Valero, E, Dequin, S, and Casal, M. 2004. Survey of molecular methods for
577 the typing of wine yeast strains. FEMS Microbiol. Lett. 231: 19-26.
578 Torija MJ, Rozes, N, Poblet, M, Guillamon, JM, and Mas, A. 2001. Yeast population
579 dynamics in spontaneous fermentations: Comparison between two different wine-
580 producing areas over a period of three years. Antonie van Leeuwenhoek 79: 345-
581 352.
582 Valero E, Schuller, D, Gambon, B, Casal, M, and Dequin, S. 2005. Dissemination and
583 survival of commercial wine yeast in the vineyard: A large-scale, three-years
584 study. FEMS Yeast Research 5: 959-969.
585 Vaughan-Martini A, and Martini, A. 1995. Facts, myths and legends on the prime
586 industrial microorganism. J. Ind. Microbiol. 14: 514-522.
587 Versavaud A, Courcoux, P, Roulland, C, Dulau, L, and Hallet, J-N. 1995. Genetic
588 diversity and geographical distribution of wild *Saccharomyces cerevisiae* strains
589 from the wine-producing area of Charentes, France. Appl. Env. Microbiol. 61:
590 3521-3529.
591 Wright S. 1978. Evolution and the genetics of populations, vol. 4. University of
592 Chicago Press, Chicago.
593
594

595 Table 1

596 Characteristics of the 6 microsatellite loci ScAAT1 – ScAAT6 that were used as genetic
597 markers in the present study.

598

599 Table 2

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

604

605 Table 3

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

608

609 Table 4

610 AMOVA analysis, F_{ST} values and distribution of variance components (%) among
611 groups (AG), among populations within groups (APWG), and within populations (WP)
612 based on microsatellite data for defined populations, including or not the genotypes of
613 commercial strains that were found in some of the sampling sites, as indicated in Figure
614 1.

615

616 Figure 1

617 Geographic location of the three vineyards A, C and P in the Vinho Verde Region, with
618 indication of the sampling sites (PI-PVI, AI-AVI and CI-CVI), the wineries (W). The

619 table summarizes the number of strains with unique genotypes for each sampling site
620 and year. The same superscript letters (a-q) represent identical genotypes in different
621 samples. Genotypes of commercial yeast strains, that were isolated from different
622 samples, are indicated by numbers (❶ Zymaflore VL1; ❷ Zymaflore F10; ❸
623 Zymaflore F15; ❹ Uvaferm BDX; ❺ ICV D254; ❻ Zymaflore VL3; ❼ Lalvin Cy
624 3079).
625
626 Figure 2
627 Alleles of microsatellite loci ScAAT1 – ScAAT6 and their frequencies in *S. cerevisiae*
628 in each of the vineyards A (light grey bars), C (dark grey bars) and P (black bars).
629 ● New alleles, identified in the present study; ▲ Alleles with major differences
630 regarding their frequency of occurrence in each vineyard; a, c, p Unique alleles,
631 occurring in only in vineyards A, C and P, respectively.
632
633 Figure 3
634 UPGMA phenogram based on Euclidean distance of allelic frequencies from strains
635 found at each sampling site over 3 years excluding (a) or including (b) the genotypes of
636 commercial yeast strains. Numbers in parenthesis indicate the number of strains
637 corresponding to unique patterns.
638
639

640 Table 1

641

Microsatellite designation	Repeat	ORF or coordinates	Chromo-some	Primers	Fluoro-chrome	Size (S288C)	Nº of repeats (S288C)
ScAAT1	ATT	86 901 – 87 129	XIII	F: AAAAGCGTAAGCAATGGTGTAGAT 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: AACCCCCATTTCTCAGTCGGA	TET	304	12
ScAAT5	TAA	897 028 - 897 259	XVI	F: GCCAAAAAAAAATAATAAAAAA R: GGACCTGAACGAAAAGAGTAG	TET	231	13
ScAAT6	TAA	105 661 – 105 926	IX	F: TTACCCCTCTGAATGAAAACG R: AGGTAGTTTAGGAAGTGAGGC	HEX	266	19

642

643

644 Table 2

Designation (Figure 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

645

646 Table 3

647

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

648

649 Table 4

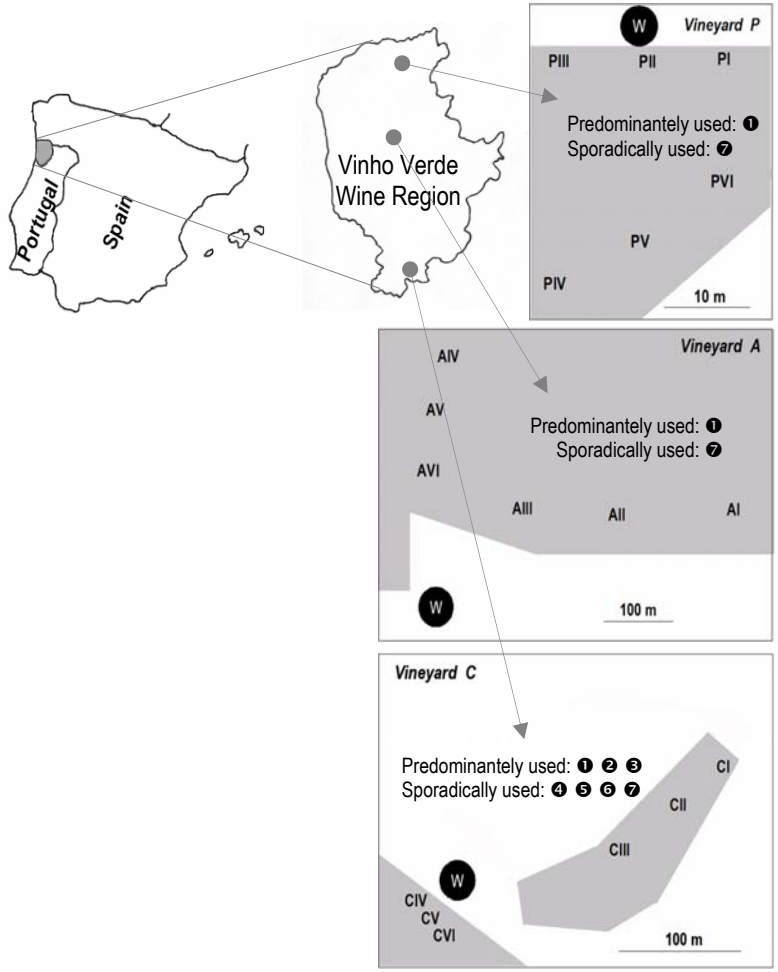
Source of variation			- commercial strains					+ commercial strains				
			AG	AGWP	WP	F_{ST}	P (r < o)	AG	AGWP	WP	F_{ST}	P (r < o)
Among vineyards	2001	A/P	3.03	9.03	87.94	0.12	< 0.0001	3.68	6.94	89.39	0.11	< 0.0001
	2002		6.38	13.28	80,33	0.20	0.0001	5.60	11.92	82.48	0.18	< 0.0001
	2003		2.76	11.29	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 + 2002 + 2003	A	-0.02	16.65	83.38	0.17	< 0.0001	0.48	15.99	83.53	0.16	< 0.0001
		C	-12.27	24.46	87.81	0.12	0.0001	-8.31	18.78	89.53	0.10	< 0.0001
		P	-1.23	9.19	92.05	0.08	< 0.0001	-0.82	6.88	93.94	0.06	0.0001

650
651

652

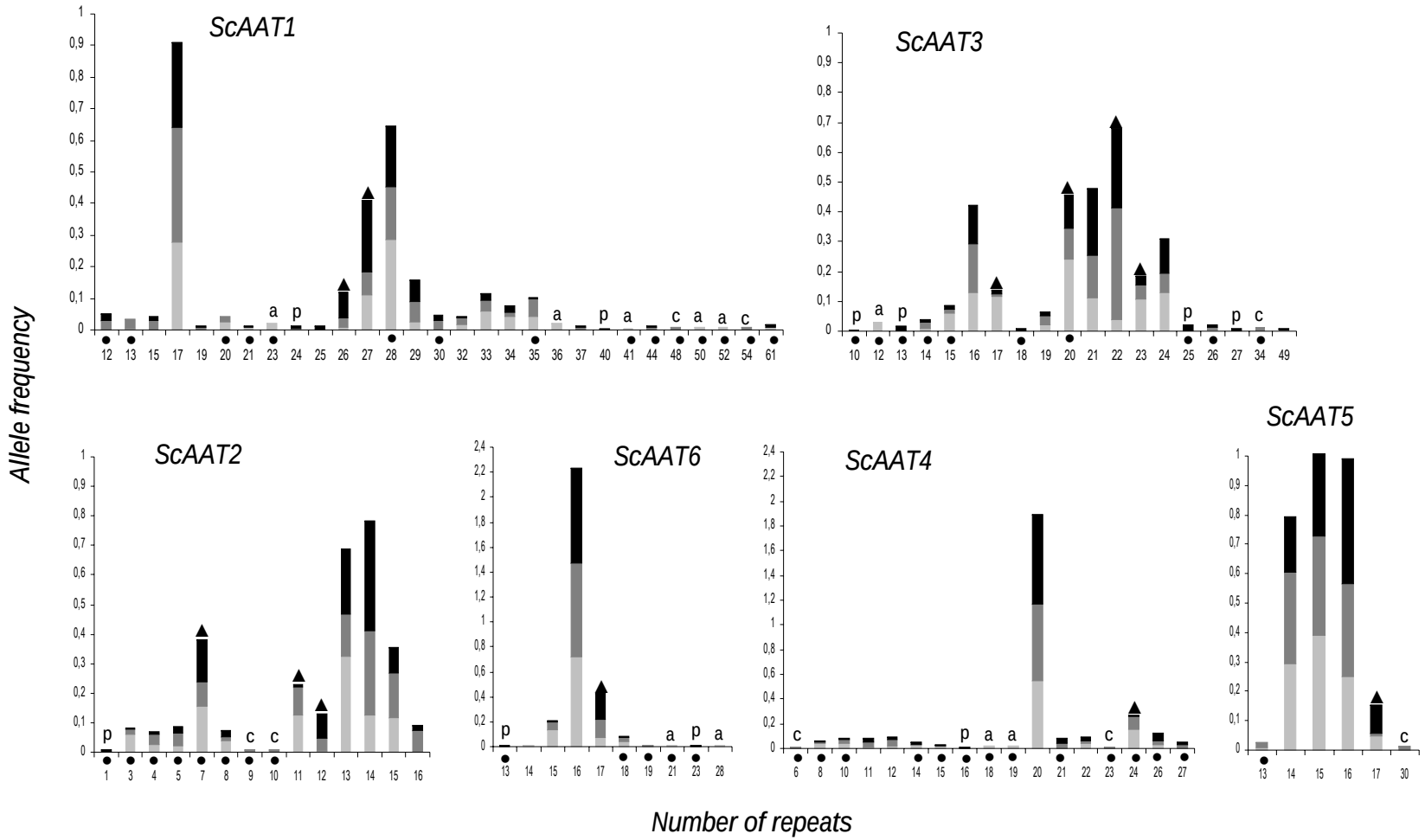
653

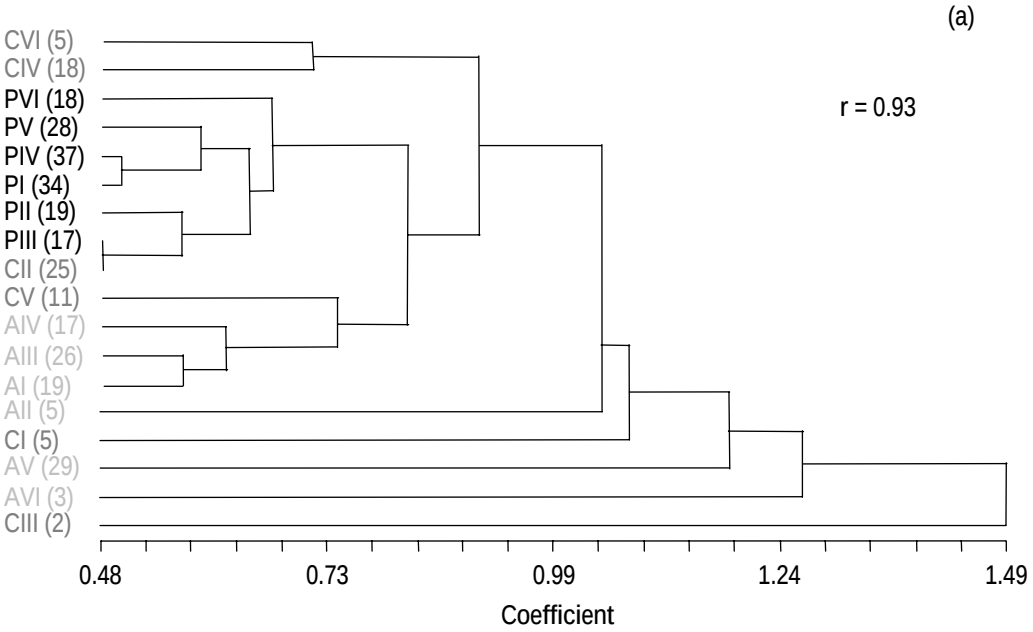
654 Figure 1

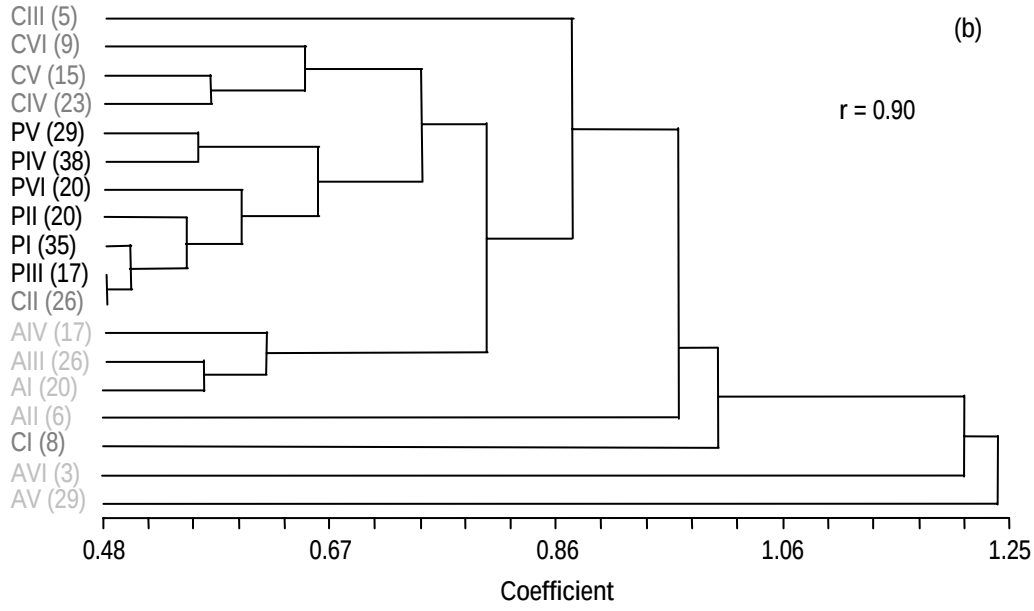


	Site					
	I	II	III	IV	V	VI
2001	7 ❶	15 b g ❶	7	20 g h i	15 h ❶	13 ❶❺
2002	3	4 b	1 b	-	10 i ❶	1 b
2003	25 e j ❶	2 e	9 b	18 k ❶❼	5	4
2001	-	2 a	8	-	-	-
2002	16 a b c	2 b	9 c d	5 a	9 d	1
2003	5 b ❶	3 b ❶	9 a	12	20	2
2001	-	20 ❺	4 ❷❸❺	5 ❶❷❸❹❺	6 ❶❷❸❺	8 ❷❸❹❺
2002	-	-	-	1	-	-
2003	8 b j ❶❷❻	6	1	18 ❸	9 k	2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

Figure 2


ScAAT1
ScAAT3
ScAAT2
ScAAT6
ScAAT4
ScAAT5
Allele frequency
Number of repeats


(a)
CVI (5)
CIV (18)
PVI (18)
PV (28)
PIV (37)
PI (34)
PII (19)
PIII (17)
CII (25)
CV (11)
AIV (17)
AIII (26)
AI (19)
AII (5)
CI (5)
AV (29)
AVI (3)
CIII (2)
r = 0.93
0.48
0.73
0.99
1.24
1.49
Coefficient


(b)
CIII (5)
CVI (9)
CV (15)
CIV (23)
PV (29)
PIV (38)
PVI (20)
PII (20)
PI (35)
PIII (17)
CII (26)
AIV (17)
AIII (26)
AI (20)
AII (6)
CI (8)
AVI (3)
AV (29)
r = 0.90
0.48
0.67
0.86
1.06
1.25
Coefficient