FEMS Yeast Research, 15, 2015, fov063

doi: 10.1093/femsyr/fov063
Advance Access Publication Date: 17 July 2015
Research Article

RESEARCH ARTICLE

Intrastrain genomic and phenotypic variability of the commercial *Saccharomyces cerevisiae* strain Zymaflore VL1 reveals microevolutionary adaptation to vineyard environments

Ricardo Franco-Duarte[1,*], Frédéric Bigey[2], Laura Carreto[3], Inês Mendes[1], Sylvie Dequin[2], Manuel AS Santos[3], Célia Pais[1] and Dorit Schuller[1]

[1]CBMA (Centre of Molecular and Environmental Biology)/Department of Biology/University of Minho, Campus de Gualtar, 4710–057 Braga, Portugal, [2]INRA, UMR1083, Sciences pour l'Oenologie, 34060 Montpellier, France and [3]RNA Biology Laboratory, CESAM, Biology Department, Aveiro University, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

*Corresponding author: CBMA (Centre of Molecular and Environmental Biology)/Department of Biology/University of Minho, Campus de Gualtar, 4710–057 Braga, Portugal. Tel: (+351) 253601522; Fax: (+351) 253-678-980; E-mail: ricardofilipeduarte@bio.uminho.pt

One sentence summary: The maintenance of microbial species in different environmental conditions is associated with adaptive microevolutionary changes that are shown here to occur within the individuals of the same strain in comparison with the reference.
Editor: Cletus Kurtzman

ABSTRACT

The maintenance of microbial species in different environmental conditions is associated with adaptive microevolutionary changes that are shown here to occur within the descendants of the same strain in comparison with the commercial reference strain. However, scarce information is available regarding changes that occur among strain descendants during their persistence in nature. Herein we evaluate genome variations among four isolates of the commercial winemaking strain *Saccharomyces cerevisiae* Zymaflore VL1 that were re-isolated from vineyards surrounding wineries where this strain was applied during several years, in comparison with the commercial reference strain. Comparative genome hybridization showed amplification of 14 genes among the recovered isolates being related with mitosis, meiosis, lysine biosynthesis, galactose and asparagine catabolism, besides 9 *Ty* elements. The occurrence of microevolutionary changes was supported by DNA sequencing that revealed 339–427 SNPs and 12–62 indels. Phenotypic screening and metabolic profiles also distinguished the recovered isolates from the reference strain. We herein show that the transition from nutrient-rich musts to nutritionally scarce natural environments induces adaptive responses and microevolutionary changes promoted by *Ty* elements and by nucleotide polymorphisms that were not detected in the reference strain.

Keywords: adaptation; *Saccharomyces cerevisiae*; next-generation sequencing; micro-evolutionary changes; phenotypic characterization

Received: 2 January 2015; **Accepted:** 10 July 2015
© FEMS 2015. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

INTRODUCTION

Saccharomyces cerevisiae strains from diverse natural habitats harbor a vast amount of phenotypic (Gasch *et al.* 2000; Kvitek, Will and Gasch 2008; Liti *et al.* 2009; Umek *et al.* 2009; Camarasa *et al.* 2011; Warringer *et al.* 2011; Mendes *et al.* 2013) and genetic diversity (Schuller *et al.* 2005; Umek *et al.* 2009; Dequin and Casaregola 2011; Franco-Duarte *et al.* 2011; Roberts and Oliver 2011; Borneman, Pretorius and Chambers 2013) driven by interactions between yeast and the respective environment. During the long history of association between *S. cerevisiae* strains and human activity, the genomic makeup of this yeast has been shaped through the action of multiple independent rounds of wild yeast domestication. Recently published results showed that the species as a whole consists of both 'domesticated' and 'wild' populations, whereby the genetic divergence is associated with both ecology and geography (Liti *et al.* 2009; Schacherer *et al.* 2009; Liti and Schacherer 2011). Sequence comparisons by low-coverage whole-genome sequencing and high-density arrays revealed few well-defined geographically isolated lineages, and many mosaic lineages, suggesting the occurrence of two domestication events during the history of association with human activities, one for sake strains and one for wine yeasts. 'Wild' populations are mostly associated with oak trees, nectars or insects (Greig and Leu 2009; Liti *et al.* 2009; Schacherer *et al.* 2009), while winemaking *S. cerevisiae* isolates form a genetically differentiated group, distinct from 'wild' strains and also from strains associated with other fermentations (sake and palm wine) or clinical strains. This is sustained by the fact that the oldest lineages and the majority of variation were found in strains from sources unrelated to wine production (Fay and Benavides 2005).

The diversifying selection imposed after yeast expansion into new environments, due to unique pressures, lead to interstrain variability (Diezmann and Dietrich 2009; Dunn *et al.* 2012; Borneman, Pretorius and Chambers 2013), resulting many times in adaptive genomic changes, such as gene amplifications, chromosomal-length variations, chromosomal rearrangements (especially amplifications and deletions) and copy-number increases (Adams *et al.* 1992; Goto-Yamamoto *et al.* 1998; Dunham *et al.* 2002; Pérez-Ortin *et al.* 2002; Carro *et al.* 2003; Schacherer *et al.* 2007; Borneman *et al.* 2008; Diezmann and Dietrich 2009; Liti *et al.* 2009; Dunn *et al.* 2012; Bleykasten-Grosshans, Friedrich and Schacherer 2013). Retrotransposons are known by their key role in the generation of genomic variability in *S. cerevisiae*, mediating chromosomal rearrangements that are bound by transposon-related sequences at the breakpoints (Dunham *et al.* 2002). *S. cerevisiae* strains contain several copies (between 2 and 30) of retrotransposons, being associated with karyotype alterations in natural and industrial strains, as reviewed in Bleykasten-Grosshans and Neuvéglise (2011).

Genomic variation among *S. cerevisiae* strains has been inferred by several methods in the past years, such as microsatellite amplification (Howell *et al.* 2004; Schuller *et al.* 2005; Legras *et al.* 2007; Schuller and Casal 2007; Muller and McCusker 2009; Richards, Goddard and Gardner 2009; Umek *et al.* 2009), comparative genome hybridization on array (aCGH) (Dunham *et al.* 2002; Winzeler *et al.* 2003; Dunn, Levine and Sherlock 2005, 2012; Carreto *et al.* 2008; Kvitek, Will and Gasch 2008) and single-nucleotide polymorphisms (SNPs) detection after sequencing (Liti *et al.* 2009; Schacherer *et al.* 2009), among others. Recent findings obtained by aCGH (Dunn, Levine and Sherlock 2005) showed copy-number amplifications, mainly in subtelomeric regions and in transposable elements among *S. cerevisiae* wine strains from different geographical origins (both commercial and from natural environments). In a similar work, the characterization of genome variability was also expanded to strains from other technological origins (Carreto *et al.* 2008). aCGH was used to detect copy-number variations in 16 yeast strains, according to their origin—laboratorial, commercial, environmental or clinical. Results showed that the absence of about one third of the *Ty* elements determined genomic differences in wine strains, in comparison to laboratorial and clinical strains, whereas subtelomeric instability related with depletions was associated with the clinical phenotype. Some of the variable genes between the analyzed groups were related with metabolic functions connected to cellular homeostasis or transport of different solutes such as ions, sugars and metals. Intrastrain differences were also revealed by Dunn *et al.* (2012) by the findings of deletions and amplifications of single genes in different isolates of the same strain that were obtained from different laboratories. In this work, the differences detected by aCGH were also extended to the phenotypic level.

With the development of 'next-generation' sequencing methods, an exponential increase was observed in the number of strains with their whole genome sequenced. In 2012, about 100 whole genome sequences of *S. cerevisiae* strains were available from different geographical and technological origins, with a large predominance of industrial strains (Borneman, Pretorius and Chambers 2013). These sequencing projects were a major breakthrough in the understanding of genomic differences among strains, mainly through the finding of numerous strain-specific open reading frames (ORFs), especially for wine strains (Argueso *et al.* 2009; Novo *et al.* 2009; Dowell *et al.* 2010; Wenger, Schwartz and Sherlock 2010; Borneman *et al.* 2011; Damon *et al.* 2011; Engel and Cherry 2013).

Within our previous work, we showed that commercial winemaking *S. cerevisiae* strains are disseminated from the wineries where they are used and can be recovered from locations in close proximity (10–200 m) (Valero *et al.* 2005). In the referred study, 100 isolates of the commercial strain Zymaflore VL1 were recovered from vineyards next to wineries where this strain was used during several years. The permanence of these isolates in natural environments induced genetic changes that were not found among a control group of isolates that derived from clonal expansion of the commercial reference strain (Schuller *et al.* 2007). These changes were mostly related with chromosomal-size variations, mainly for smaller chromosomes, loss of heterozygosity, microsatellite expansion and differences in the interdelta sequence amplification patterns. Besides, the fermentative capacity of some isolates was affected, pointing to a possible adaptive mechanism induced by genetic changes. The objective of the present work was to undertake a deeper genomic characterization of recovered isolates of the commercial strain Zymaflore VL1, using aCGH and SNP analysis, in order to investigate the extent of variation to which natural isolates differ from the reference strain. In addition, we performed an extensive phenotypic analysis using both enological and taxonomic tests, including also metabolic profiling (HPLC and GC-MS) of a must fermented by these isolates.

MATERIALS AND METHODS

Strain isolates

Strain Zymaflore VL1 is a non-indigenous diploid yeast that was originally isolated from the region of Gironde, France. Four natural isolates (VL1–018, VL1–020, VL1–099 and VL1–108) were

obtained in our previous work (Schuller *et al.* 2007) from spontaneous fermentations of grape samples collected from two different vineyards, located close to wineries where this strain had been used for winemaking in consecutive years. These isogenic isolates showed identical mitochondrial DNA restriction fragment length polymorphisms, although with small differences regarding their karyotype, microsatellite allele sizes and interdelta sequence amplification patterns. The DNA content of these isolates was identical to the reference strain, as determined by flow cytometric analysis (data not shown). The original commercial VL1 reference strain, kindly provided by *Laffort Oenologie*, was used as a reference.

DNA isolation

After cultivation of a frozen (−80°C, 30% v/v glycerol) aliquot of yeast cells in 1 mL YPD medium (yeast extract 1% w/v, peptone 1% w/v, glucose 2% w/v) during 36 h at 28°C (160 rpm), DNA isolation was performed as previously described (Schuller *et al.* 2004). DNA was then quantified (Nanodrop ND-1000) and used for aCGH and for DNA sequencing.

Comparative genome hybridization on array

For aCGH experiments, DNA microarrays were produced as referred in Carreto *et al.* (2008), being the array design and spotting protocol deposited in the ArrayExpress database under the accession code A-MEXP-1185. The labeling protocol was also performed as referred (Carreto *et al.* 2008), whereby ULS-Cy3 labeled DNA of each of the four isolates (VL1–018, VL1–020, VL1–099 and VL1–108) was combined with ULS-Cy5 labeled DNA from the commercial reference strain. Dye-swap hybridizations were performed for each isolate, ruling out potential bias introduced by inherent differences in dye incorporation. To ensure microarray data baseline robustness, differentially labeled DNA from the S288c strain were cohybridized, in a total of six self–self experiments, and used as controls. Images were obtained using Agilent G2565AA microarray scanner, and the fluorescence was quantified by image analysis using QuantArray software (PerkinElmer). Data were analyzed with BRB-ArrayTools v3.4, using median normalization. The relative hybridization signal of each ORF was derived from the average of the two dye-swap hybridizations, and deviations from the 1:1 normalized $\log_2$ ratio were taken as indicative of changes in DNA copy number. The significance of these changes was evaluated using multiclass significance analysis (SAM) and hierarchical clustering, as implemented in the TM4 software (MeV). SAM analysis (Tusher, Tibshirani and Chu 2001) indicated the ORFs with significant copy-number alteration in at least one of the strains, with an FDR of 0.336.

DNA sequencing and SNP detection

Genomic DNA of the five isolates was processed to be sequenced according to the manufacturer´s protocols (Only *et al.* 2009), in paired-end 104 bp mode, and sequenced using an Illumina HiSeq2000 analyzer. Samples were tagged and multiplexed using a custom barcode of 6 bp length. All de-multiplexed reads were aligned to the *sacCer3* assembly of the yeast reference genome using BWA (bwa-aln and bwa-sampe; version 0.7.5a) with default parameters. Sequences were aligned using SAMtools (version 1.1) using the commands *view, sort, index and mpileup* (Li and Durbin 2009). All possible variants including frameshift insertions/deletions (indels) and SNPs were then called from the aligned sequences, using Annovar (Wang, Li and Hakonarson 2010), with the following filters: QUAL ≥ 30 (‘phred-scaled quality score for the assertion made in the alternate allele’), DP ≥ 15 (‘raw read depth’ or ‘coverage’), MQ ≥ 40 (‘root-mean-squared mapping quality of coverage reads’) and GQ ≥ 50 (‘genotype quality or phred-scaled confidence that the true genotype is the one provided’).

Phenotypic characterization

Phenotypic screening was performed considering a wide range of physiological traits that are also important from an oenological point of view, considering a previously established experimental design that included evaluation of growth by (i) measurement of optical density (A_{640}) after 22 h of growth in 96-well microplates containing white grape must plus the compound under analysis, or (ii) visual evaluation of growth in solid YPD with the compound to be tested (Mendes *et al.* 2013). Thirty phenotypic tests were considered, as shown in Table 3, and all results were assigned to a class between 0 and 3 [0: no growth ($A_{640} = 0.1$) or no visible growth on solid media or no color change of the BiGGY medium; 3: at least 1.5-fold increase of A_{640}, extensive growth on solid media or a dark brown colony formed in the BiGGY medium; scores 1 and 2 corresponded to the respective intermediate values].

Fermentation media and conditions

Triplicate fermentations (18°C, 150 rpm) of each of the five isolates were carried out with grape must of the variety Loureiro, using Erlenmeyer flasks (100 mL) with rubber stoppers that were perforated with a syringe needle for CO_2 release. The fermentative progress of each isolate was recorded by weight loss determination due to CO_2 liberation. Samples were collected and frozen (−20°C) for metabolic analysis when fermentation ended (constant weight, when no more CO_2 was released).

Bioanalytical methods

High-performance liquid chromatography with refractive index (HPLC-RI) was used to quantify fructose, glucose, ethanol, glycerol and organic acids (malic, acetic and succinic). Prior to analysis, supernatant samples were filtered through a 0.22-μm pore filter, and then analyzed in an EX Chrome Elite HPLC, using a Rezex® Ion Exclusion column. Column and refractive index detector temperatures were 60 and 40°C, respectively, and the flow rate was 0.50 mL min^{-1} from 0 to 9 min and from 15 to 35 min of run length and 0.25 mL min^{-1} from 10 to 14 min.

Higher alcohols, esters and fatty acids were determined by GC-MS. Analyses were performed by solid phase microextraction (SPME), using a divinylbenzene/carboxen/polydimethylsiloxane fiber, and 4-methyl-2-pentanol as internal standard. Samples were analyzed using a Thermo-Finnigan Trace-GC with a single Quadrupole Trace-DSQ Mass Selective Detector (Thermo Electron Corporation, USA), equipped with a Zebron ZB-FFP capillary column. The injector temperature was set to 260°C and the flow rate to 0.8 mL min^{-1}, with helium used as the carrier gas. GC-MS concentrations of volatile compounds were normalized using maximum normalization, and differences between the isolates were represented using principal component analysis (PCA) of the Unscrambler X software (Camo Inc.).

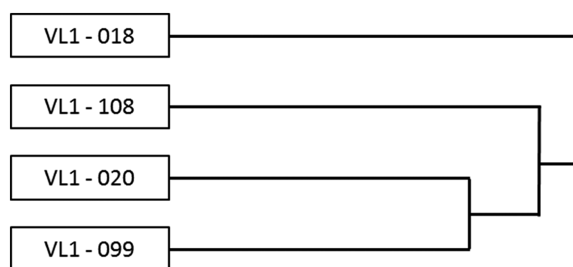

Figure 1. Hierarchical clustering of the aCGH profiles. All of the four natural isolates and the isolate derived from the commercial reference strain were used in the hierarchical clustering analysis, using Pearson correlation with average linkage of the normalized aCGH data.

Statistical analysis

Statistical analyses were performed with the data set obtained from HPLC quantification, using two-sample paired *t*-test, comparing always each set of data with the reference strain data set, and considering results in which $P < 0.05$ as significant.

RESULTS

Genomic changes revealed by aCGH profiles

Comparative genome analysis of the isolates was conducted using microarrays containing 70 mer probes designed from the genome sequence of strain S288c, targeting 6388 ORFs. Genomic DNA of the recovered isolates of the commercial winemaking strain *S. cerevisiae* Zymaflore VL1 (VL1–018, VL1–020, VL1–099 and VL1–108) was fluorescently labeled and competitively hybridized with the DNA of the VL1 reference strain. Hybridizations were performed in duplicate, in reverse Cy-dye labeling (dye-swap) design (see the section 'Methods'). Figure 1 shows the global genome variability of the hierarchical cluster analysis of the aCGH data. Isolate VL1–018 was most differentiated from the remaining isolates that grouped into two clusters (VL1–108 and VL1–020/VL1–099).

Multiclass significance analysis (SAM, MeV software) was used to evaluate genomic changes between reisolated yeasts and the reference strain using S288c chromosomal coordinates. ORF copy-number alterations occurred in all four recovered isolates, in comparison with the VL1 reference strain (Table 1; Fig. 2). All genomic alterations corresponded to copy-number amplifications, whereas deletions were not detected. The 22 amplified ORFs showed a stochastic distribution among 10 chromosomes, so that each of the recovered isolates had a unique amplification pattern. As summarized in Table 1, copy-number increases (between 1.5- and 2.0-fold) were associated with 14 annotated ORFs in isolates VL1–020 and VL1–099, mainly related with mitosis (*SHE1*), meiosis (*HFM1*), lysine biosynthesis (*LYS14*), galactose (*GAL1*) and asparagine catabolism (*ASP3–2*). *ASP3–2* amplification might be a response to nitrogen starvation (Bon *et al.* 1997), whereas *GAL1* amplification, which is expressed in the beginning of the galactose catabolism, might be important for the improved use of galactose as alternative carbon source. Nine ORFs with increased copy numbers (between 1.5- and 1.8-fold) corresponded to amplified *Ty* elements, in isolates VL1–018 (1), VL1–099 (2) and VL1–108 (6).

Sequence analysis of isolates recovered from vineyards

To investigate the extent of variation to which natural isolates differ from the reference strain, we sequenced DNA from the recovered isolates and from the reference strain by Illumina sequencing. Short sequence reads (104 bp) were processed and aligned to the reference genome of strain S288c using BWA and SAMtools. Functional annotation of genetic variants between each of the tested genomes and the reference VL1 strain were called using ANNOVAR. Quantification of SNPs and indels was performed by comparison of each recovered isolate with the reference strain using some filters to control false positives (Table S1, Supporting Information). From the initial 2610 variants called between the natural isolates and the reference strain, 1144 did not pass the quality filters. MQ and GQ were the ones eliminating more false variants. This filtration was found to be imperative to eliminate artifacts that were initially considered as true variants by the method.

Our results (Table 2) show that intrastrain differences between natural isolates and the VL1 reference strain were in the range of 339–427 SNPs called for strains VL1–018 and VL1–108, respectively. VL1 intrastrain variation of recovered isolates revealed between 12–62 insertions and deletions (indels). From the 1466 total variants called (after filtration), 958 corresponded to changes in coding regions (~65%). Regarding SNPs, the large majority led to changes in the genotype from homozygous to heterozygous, as expected in diploid strains. Also, 32% of SNPs called (433 out of 1351) corresponded to non-synonymous mutations.

The distribution of SNPs and indels per chromosome in the natural isolates is shown in Fig. 3. The majority of SNPs that lead to a change of genotype from homozygous to heterozygous was detected in chromosome II (panel A), being almost residual in the remaining chromosomes. Regarding heterozygous–homozygous changes of genotypes, they were located mainly in chromosome IX (20 to 47 SNPs identified) being similarly distributed in the remaining chromosomes. Indels showed also to be a predominance in chromosome II, with a maximum of 11 insertions and deletions identified for three of the isolates in this chromosome (panel C).

Phenotypic characterization

To evaluate the extent of phenotypic variation, a screening approach was devised, taking into consideration 30 phenotypic tests, including also tests that are important for winemaking strain selection. High-throughput testing in microplates was performed using supplemented grape must, and optical density (A_{640}) was measured after 22 h of incubation. Growth in solid culture media (BiGGY medium, Malt Extract Agar supplemented with ethanol and sodium metabisulphite) was evaluated by visual scoring. The patterns of phenotypic variation are summarized in Table 3. Fourteen phenotypic traits distinguished the group of the four recovered isolates from the reference strain, which was unable to grow at 18°C, but evidenced some growth in the presence of $CuSO_4$ (5 mM) and SDS 0.01% (v/v). Variable growth patterns were found between some of the natural isolates in relation to the reference strain, regarding KCl (0.75 M), NaCl (1.5 M), $KHSO_3$ (300 mg L^{-1}), wine supplemented with glucose (0.5% and 1% w/v), ethanol (14, 16 and 18% w/v) + $Na_2S_2O_5$, cycloheximide (0.05 and 1 μg mL^{-1}) and galactosidase activity. The main differences were observed between the natural isolates and the reference strain, whereas small changes were observed among the four natural isolates, for example in terms of ethanol resistance. For the analyzed tests, phenotypic differences were limited to the transition from one phenotypic class to another, and also presented a stochastic distribution of

Table 1. Genes with amplified copy-number changes, as detected by SAM analysis of aCGH data.

Strain	Systematic name	Classical name	SGD	Chromosome	Fold change
VL1 – 018	YMR046C	–	Ty element	13	1.7
VL1 – 020	YBL031W	SHE1	Mitotic spindle protein	2	1.7
	YOR019W	NA	Unknown function; may interact with ribosomes	15	1.9
	YGL251C	HFM1/MER3	Meiosis-specific DNA helicase involved in the conversion of double-stranded breaks	7	1.5
	YOR155C	ISN1	Catalyzes the breakdown of inosine 5′ monophosphate to inosine	15	1.6
	YDR034C	LYS14	Transcriptional activator involved in regulation of genes of the lysine biosynthesis pathway	4	1.7
	YBR020W	GAL1	Phosphorylates alpha-D-galactose to alpha-D-galactose-1-phosphate in the first step of galactose catabolism	2	1.9
VL1 - 099	YDR120C	TRM1	tRNA methyltransferase	4	1.6
	YLR407W	NA	Unknown function	12	1.7
	YOR260W	GCD1/TRA3	Gamma subunit of the translation initiation factor eIF2B	15	1.7
	YKL102C	NA	Dubious ORF unlikely to encode a functional protein; deletion confers sensitivity to citric acid	11	1.6
	YOR257W	CDC31/DSK1	Calcium-binding component of the spindle pole body half-bridge; binds multiubiquitinated proteins and is involved in proteasomal protein degradation	15	1.7
	YHR212C	NA	Dubious ORF; unlikely to encode a functional protein	8	1.7
	YLR157C	ASP3–2	Cell-wall L-asparaginase II involved in asparagine catabolism; expression induced during nitrogen starvation	12	1.7
	YPL218W	SAR1	GTPase, GTP-binding protein of the ARF family; required for transport vesicle formation during ER to Golgi protein transport	16	2.0
	YHL009W-A	–	Ty element	8	1,6
	YHL009W-B	–	*Ty element*		1.6
VL1 - 108	YBL005W-A	–	*Ty element*	2	1.5
	YDR170W-A	–	*Ty element*	4	1.7
	YDR210C-C	–	*Ty element*	4	1.5
	YGR161C-C	–	*Ty element*	7	1.5
	YMR046C	–	*Ty element*	13	1.7
	YNL284C-A	–	*Ty element*	14	1.8

variation among the isolates, as previously observed for aCGH results and sequence analysis.

Fermentative profiles and metabolic characterization

Triplicate fermentations were carried out with each of the five isolates, using white grape must. HPLC and GC-MS analysis were performed with samples obtained from the end of fermentation (at constant weight, when no more CO_2 was released) to evaluate the chemical compounds associated with the differences described in the previous sections. A very good reproducibility regarding fermentation profile and time was obtained among the three fermentation replicates with the exception of a small delay in the maximum CO_2 release for isolate VL1–099 (Fig. 4).

Strain-dependent differences could be observed regarding final concentrations of organic acids (malic, succinic and acetic), fructose, glycerol and ethanol (Fig. 5). Malic acid concentration ranged, for all the isolates, between 6.3 and 7.1 g L^{-1}, whereas acetic and succinic acids ranged between 0.57–0.65 g L^{-1} and 0.39–0.44 g L^{-1}, respectively. Final concentrations of ethanol, glycerol and fructose ranged between 159–175, 6.64–8.07 and 5.4–35.2 g L^{-1}, respectively. Statistical significance ($P \leq 0.01$ for two-paired sample *t*-test) was obtained only for the isolates VL1–099 and VL1–108, regarding the concentrations of malic acid, ethanol and/or glycerol. Another compound that explained variability between isolates was fructose, although not in a statistically significant way. This sugar was still present in values around 30 g L^{-1}, indicating that these isolates do not assimilate fructose in large amounts. Isolate VL1–108 produced higher amounts of ethanol and showed a more reduced fructose concentration.

Aromatic compounds from the final fermentation stage were quantified by GC-MS after SPME. PCA of the GC-MS data (Fig. 6) shows the segregation of the six isolates (scores; panel A) and the loadings for aromatic compounds (panel B) in the first two PCA components, which explain 91% of the observed variability between isolates. The consideration of further components

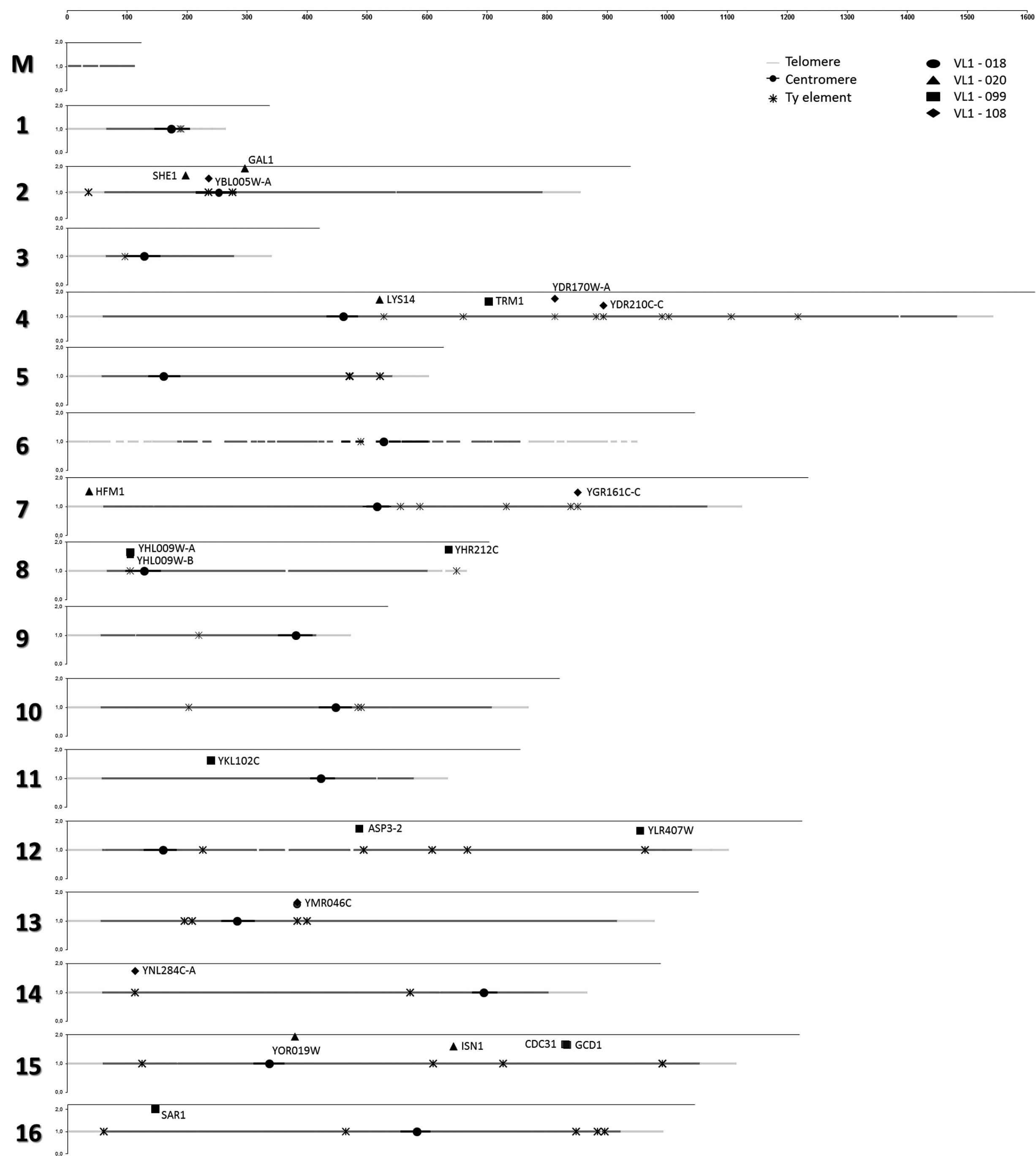

Figure 2. Graphical representation of gene copy-number alterations for the 17 chromosomes (from I to XVI; plus mitochondrial DNA- M) of natural isolates, in comparison to the original reference strain, obtained by SAM analysis of aCGH data. Using annotated ORF coordinates of strain S288C, global chromosome plots are shown, indicating also ORFs with copy-number changes, as detected by SAM analysis of aCGH data. For each chromosome, the telomere and the centromere are marked, together with the locations of the *Ty* elements (relative to the S288C genome). Fold-change alterations in terms of copy number are represented by the distance of the gray symbols to the basal line, for each of the natural isolates, in comparison to the reference strain.

did not improved the explanation of variability. Panel A shows that the global aromatic profile of isolates VL1–108 and VL1–018 was very similar and most different from the reference strain. Isolate VL1–099 was the one with more similarities to the reference strain, due to its position in the PCA plot. These differences can be explained by some of the loadings of panel B, which have the most discrimination power due to their position far from the center of coordinates, namely benzene ethanol (AI), 2-methyl-1-butanol (J) and isobutanol (G), followed by ethyl lactate (P) in a smaller extent.

DISCUSSION

Saccharomyces cerevisiae has been used for a long time as a model to study responses to environmental stress. Changed environmental conditions require an efficient adaptation, mediated

Table 2. Number of nucleotide variants (SNPs and Indels) in comparison with the reference VL1 strain (Hom—homozygous; Het—heterozygous; SNPs—single nucleotide polymorphisms; Indels—insertions and deletions in the genome).

Strain	Total number of variants called	Functional impact (variants in coding regions)	Number of SNPs Total	Hom to Het change of genotype	Het to Hom change of genotype	Non-synonymous	Indels
VL1 - 018	339	228	277	5	272	123	62
VL1 – 020	341	226	324	298	26	102	17
VL1 – 099	359	245	347	205	142	106	12
VL1 - 108	427	259	403	242	161	102	24
Total	1466	958	1351	750	601	433	115

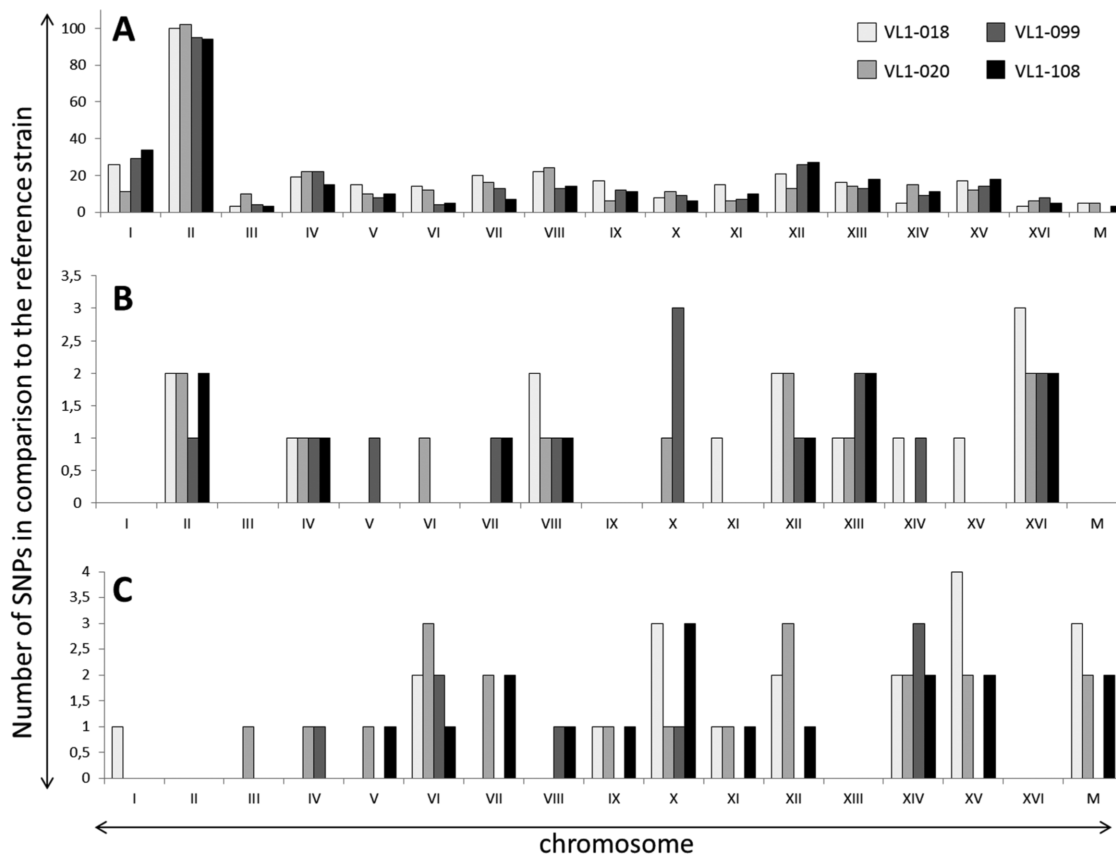

Figure 3. Number of frameshift SNPs and Indels per chromosome in the natural isolates, in comparison to the reference strain. **A**—SNPs (homozygous to heterozygous change of genotype); **B**—SNPs (heterozygous to homozygous change of genotype); **C**—Indels.

by changed gene expression to maintain cellular homeostasis. Yeast strains cultivated for longer periods under specific conditions present chromosomal rearrangements, chromosomal-length variations or other genomic changes such as gene amplifications or copy-number changes (Rachidi, Barre and Blondin 1999; Dunham *et al.* 2002; Brion *et al.* 2013). These alterations, being either neutral, beneficial or detrimental, are known to lead to phenotypic diversity, as reviewed by Bisson (2012). The loss of one or two copies of a gene can be compensated by the level of expression of the remaining copy, or by the amplification of a homolog from another chromosome. Another contributing factor is the mobile *Ty* elements in *S. cerevisiae* that can be excised and inserted along the genome, which leads to phenotypic diversity when inserted into a gene or a regulatory region.

In the present study, four isolates of the commercial strain *S. cerevisiae* Zymaflore VL1 were used, that were recovered from the environment of two vineyards that are located in close proximity to the wine cellars where this commercial yeast was used in large quantities for at least five years. The commercial strain Zymaflore VL1 was initially isolated from a French wine region. These strains were characterized for genomic changes such as gene amplifications/deletions, and sequence analysis. aCGH results showed amplification of 14 ORFs, corresponding 10 of them to annotated ORFs (Fig. 2 and Table 1). The main functions of the amplified genes were related with mitosis (*SHE1*), meiosis (*HFM1*), lysine biosynthesis (*LYS14*), galactose (*GAL1*) and asparagine catabolism (*ASP3–2*). The existence of additional copies of *GAL1* in natural isolates indicates adaptation to an environment with less amounts of glucose. In nature, galactose occurs

Table 3. Phenotypic classes regarding values of optical density (Class 0: $A_{640} = 0.1$; Class 1: $0.2 < A_{640} > 0.4$; Class 2: $0.5 < A_{640} > 1.0$; Class 3: $A_{640} > 1.0$), growth patterns in solid media or colour change in BiGGY medium, for 30 phenotypic tests.

Phenotypic test	Type of medium	VL1 - 018	VL1 - 020	VL1 - 099	VL1 - 108	VL1 - reference
30°C	Liquid (must)	3	3	3	3	3
18°C	Liquid (must)	**1**	**1**	**1**	**1**	0
40°C	Liquid (must)	3	3	3	3	3
pH 2	Liquid (must)	0	0	0	0	0
pH 8	Liquid (must)	2	2	2	2	2
KCl (0.75 M)	Liquid (must)	2	**3**	2	2	2
NaCl (1.5 M)	Liquid (must)	1	1	1	**0**	1
$CuSO_4$ (5 mM)	Liquid (must)	**0**	**0**	**0**	**0**	1
SDS (0.01% w/v)	Liquid (must)	**0**	**0**	**0**	**0**	1
Ethanol 6% (v/v)	Liquid (must)	3	3	3	3	3
Ethanol 10% (v/v)	Liquid (must)	2	2	2	2	2
Ethanol 14% (v/v)	Liquid (must)	1	1	1	1	1
Ethanol 12% (v/v)	Solid (MEA)	2	2	2	2	2
Ethanol 12% (v/v) + $Na_2S_2O_5$ (75 mg/L)	Solid (MEA)	3	3	3	3	3
Ethanol 12% (v/v) + $Na_2S_2O_5$ (100 mg/L)	Solid (MEA)	0	0	0	0	0
Ethanol 14% (v/v) + $Na_2S_2O_5$ (50 mg/L)	Solid (MEA)	3	3	**2**	3	3
Ethanol 16% (v/v) + $Na_2S_2O_5$ (50 mg/L)	Solid (MEA)	3	3	**2**	3	3
Ethanol 18% (v/v) + $Na_2S_2O_5$ (50 mg L^{-1})	Solid (MEA)	1	1	1	**2**	1
$KHSO_3$ (150 mg L^{-1})	Liquid (must)	3	3	3	3	3
$KHSO_3$ (300 mg L^{-1})	Liquid (must)	**1**	**1**	2	2	2
Wine supplemented with glucose (0.5% w/v)	Liquid	**1**	**1**	0	0	0
Wine supplemented with glucose (1% w/v)	Liquid	1	1	**0**	**0**	1
Iprodion (0.05 mg mL^{-1})	Liquid (must)	3	3	3	3	3
Iprodion (0.1 mg mL^{-1})	Liquid (must)	3	3	3	3	3
Procymidon (0.05 mg mL^{-1})	Liquid (must)	3	3	3	3	3
Procymidon (0.1 mg mL^{-1})	Liquid (must)	3	3	3	3	3
Cycloheximide (0.05 μg mL^{-1})	Liquid (must)	**1**	2	2	**1**	2
Cycloheximide (0.1 μg mL^{-1})	Liquid (must)	**1**	**1**	**1**	**1**	2
H_2S production	Solid (BiGGY)	2	2	2	2	2
Galactosidase activity	Liquid (YNB)	**1**	**2**	3	3	3

Highlighted cells indicate the differences observed between the isolates for the mentioned test.

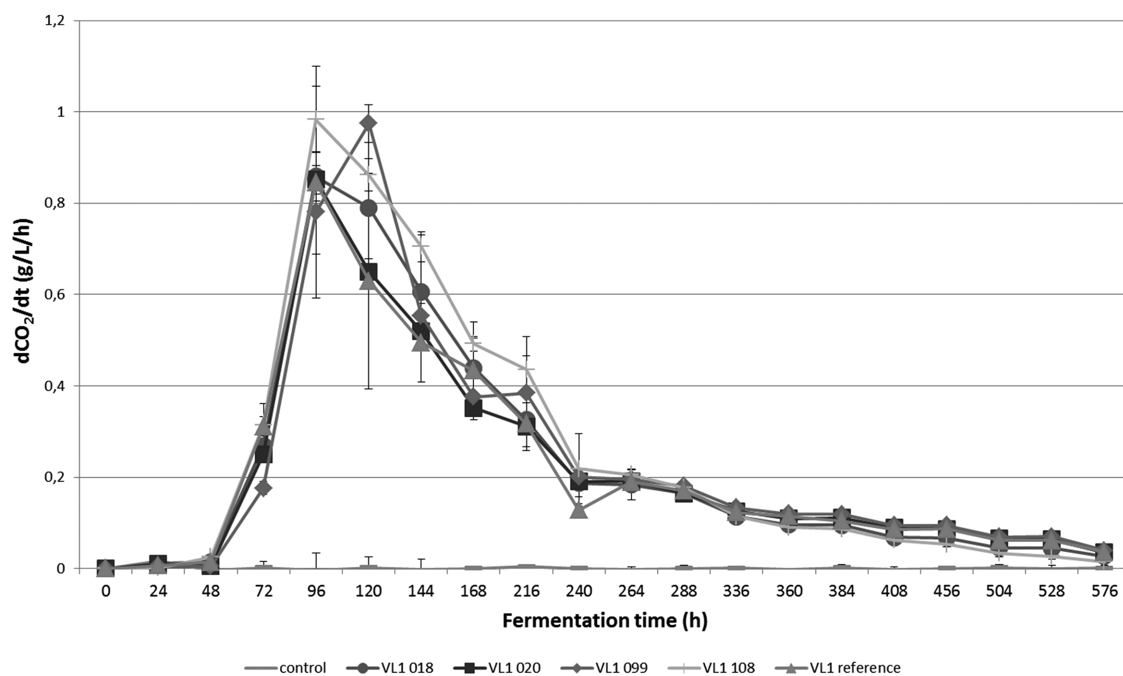

Figure 4. Fermentation profiles of four natural isolates, in comparison with the original reference strain. Values were averaged from three biological replicates ± standard deviation. Fermentations were carried out at 18°C (150 rpm) using white grape must.

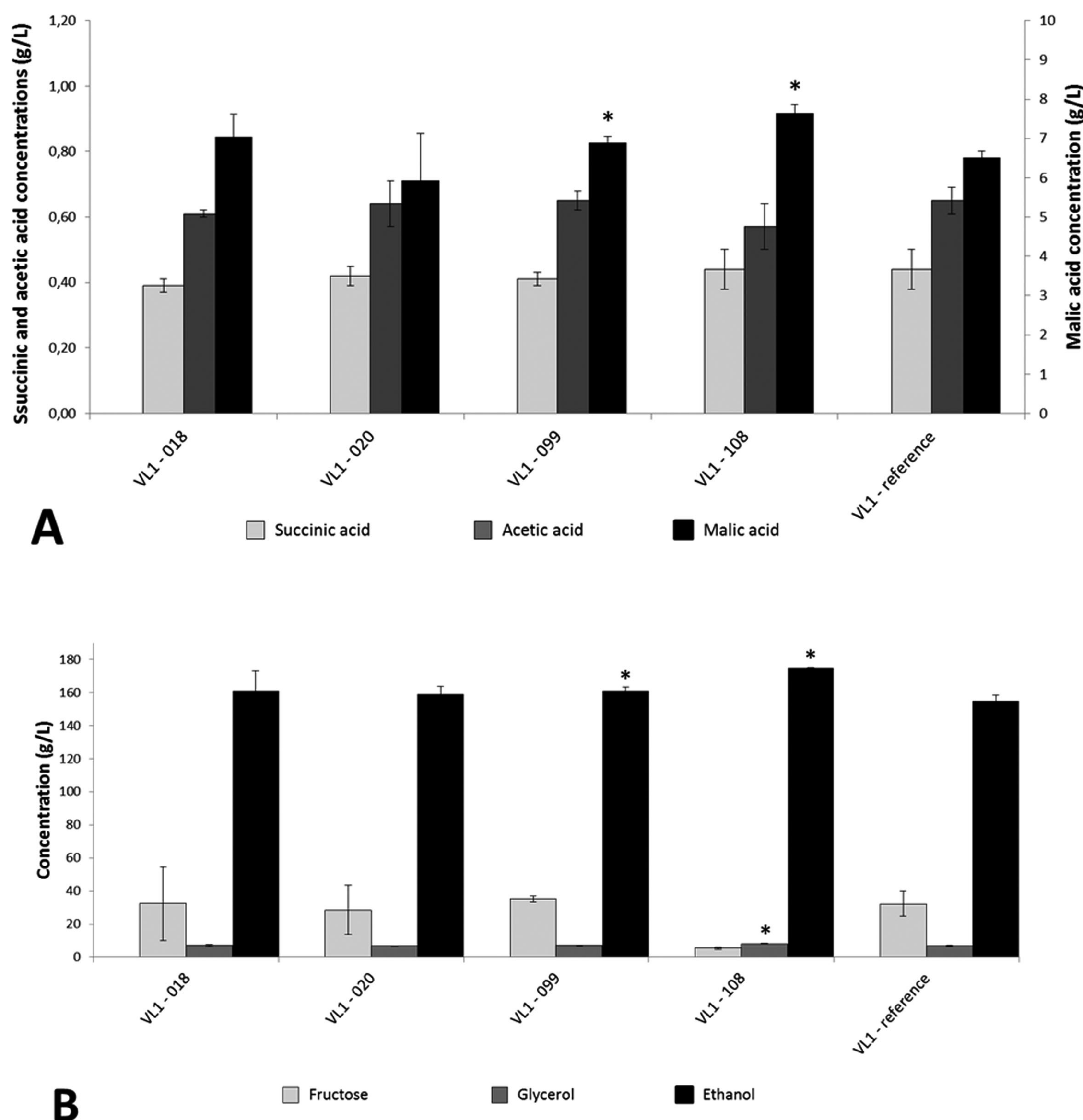

Figure 5. Concentration of (**A**) succinic, acetic and malic acids, (**B**) fructose, glycerol and ethanol, from the end of fermentations performed with natural and control isolates. Values were averaged from three biological replicates ± standard deviation, and refer to extracellular metabolites in the fermented must. Fermentations were carried out at 18°C (150 rpm) using white grape must. Statistical significance was determined using two-sample paired *t*-test. The symbol * indicates statistical significance as related to control ($P < 0.05$).

by hydrolysis of Galactan, a polymer found in hemicellulose. In *S. cerevisiae*, galactose metabolism genes are induced in the presence of galactose (Gasch *et al.* 2000), and absence of glucose (Adams 1972) underlying also glucose repression (Johnston 1999). The derepression of galactose metabolism genes in environments without glucose available has been previously described in detail (Matsumoto and Oshima 1981; St John and Davis 1981; Yocum *et al.* 1984). In the reference VL1 strain, no amplification of *GAL1* was identified, due to high glucose concentrations in the media used for the production of commercial yeasts. Copy-number amplification of gene *ASP3–2* is in agreement with the previously shown increased expression during nitrogen starvation (Jones and Mortimer 1973). These changes suggest that the recovered isolates could use asparagine as alternative nitrogen source during their presence in nature. Variable copy number of this gene was shown previously to be specific of *S. cerevisiae*, mainly from laboratory and industrial origins, being absent in other 128 fungal species (League, Slot and Rokas 2012). *ASP3–2* and four of the amplified *Ty* elements (YBL005W-A, YDR210C-C, YGR161C-C, YHL009W-A) showed also copy-number amplifications in other wine strains (Carreto *et al.* 2008). Results obtained in the mentioned study showed that the amplification of several *Ty* elements was characteristic for wine strains, contrarily to the clinical strains. The amplification of these transposable elements strengthened the importance of retrotransposition in yeast adaptation, since *Ty* sequences play a role in fragments mobilization throughout the genome.

To obtain a thorough understanding of the genomic differences between natural isolates and the reference strain, we sequenced the respective genomes and quantified SNPs and Indels (Table 2 and Fig. 3). Several studies point to the existence of several thousands of SNPs among *S. cerevisiae* strains, mainly between isolates from different technological origins. VL1

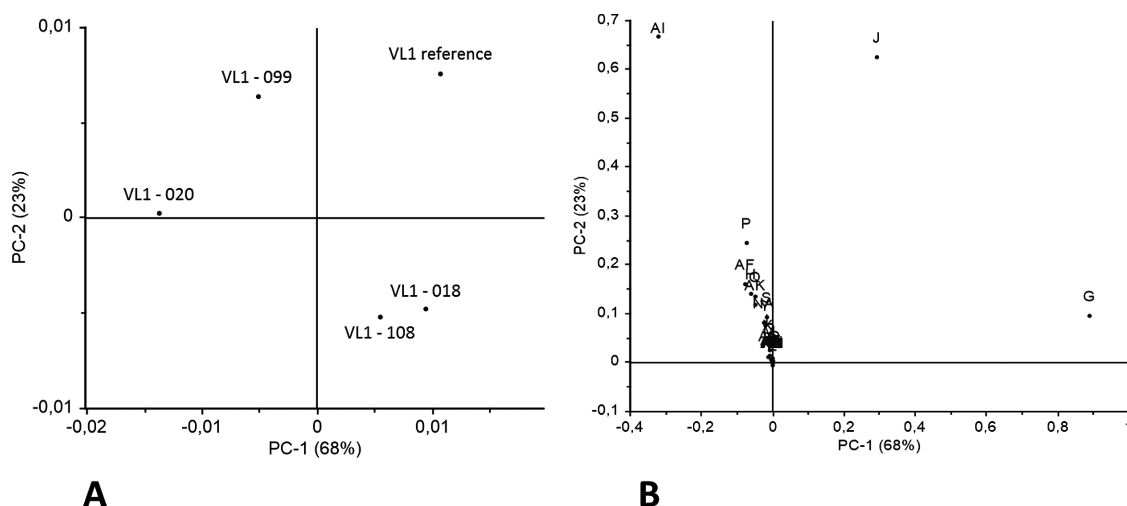

Figure 6. PCA of GC-MS data for the five isolates. Values were averaged from two biological replicates, and refer to extracellular metabolite present in the must at the end of fermentations that were carried out at 18°C (150 rpm) using white grape must. **A**—five *S. cerevisiae* isolates analyzed by GC-MS (scores); the image was zoomed-in in order to better clarify scores positioning. **B**—concentration of 41 volatile compounds determined by GC-MS (loadings). Letters indicate the following compounds: (A) dimethyl sulphide; (B) ethyl isobutyrate; (C) propyl acetate; (D) isobutyl acetate; (E) ethyl butyrate; (F) ethyl 2-methylbutyrate; (G) isobutanol; (H) isoamyl acetate; (I) methyl hexanoate; (J) 2-methyl-1-butanol; (K) 3-methyl-1-butanol; (L) ethyl hexanoate; (M) hexyl acetate; (N) ethyl heptanoate; (O) ethyl trans-2-hexenoate; (P) ethyl lactate; (Q) hexanol; (R) methyl octanoate; (S) ethyl octanoate; (T) isoamyl hexanoate; (U) octyl acetate; (V) ethyl nonanoate; (W) methyl decanoate; (X) butyric acid; (Y) ethyl decanoate; (Z) isovaleric acid; (AA) diethyl succinate; (AB) ethyl phenylacetate; (AC) 2,4,6-trichloro anisole; (AD) phenylethyl acetate; (AE) ethyl dodecanoate; (AF) hexanoic acid; (AG) guaiacol; (AH) ethyl dihydrocinnamate; (AI) benzene ethanol; (AJ) ethyl guaiacol; (AK) octanoic acid; (AL) ethyl cinnamate; (AM) 4-ethyl phenol; (AN) decanoic acid; (AO) dodecanoic acid.

reference strain showed to be different from the laboratorial strain S288c in 44 515 SNPs and 641 Indels (data not shown) being in accordance with the differences found for other strains, such as CEN.PK113–7D (21 899 SNPs and 420 Indels) (Nijkamp *et al.* 2012), YJM789 (60 000 SNPs and 6000 Indels) (Wei *et al.* 2007), M22 (5621 SNPs and 499 Indels) and YPS163 (10 773 SNPs and 423 Indels) (Doniger *et al.* 2008). However, intrastrain differences between the group of isolates obtained from nature and the reference strain consisted in just a few hundreds of SNPs, and a maximum of 33 Indels per isolate. Wine strains form a phylogenetic distinct group, some strain-specific differences, mainly in the form of insertions, were reported to be predominant in many wine strains—EC1118 (Novo *et al.* 2009), QA23, AWRI796, VL13, VIN13, FostersB, FostersO, RM11 (Borneman *et al.* 2011) and Kyokai 7 (Akao *et al.* 2011), being absent in the laboratorial strain S288c, and were related with winemaking traits (Galeote *et al.* 2010). In our reference strain—Zymaflore VL1—we detected a total of 111 unique SNPs and 8 Indels that were not detected in strain S288c, confirming the divergence of this strain, by the introduction of several changes in the laboratorial isolates in comparison with the original S288c strain. The identified isolate-specific Indels corresponded to two frameshift insertions (in chromosome II and IV), and to six frameshift deletions (chromosomes II, VIII, XII and XV). Regarding the comparison between natural isolates and the reference strain, the highest number of SNPs and frameshift insertions were detected in chromosome II, with a stochastic distribution among all natural isolates. Amplifications in this chromosome are not frequently reported in *S. cerevisiae* strains, with the exception of strain Fosters O, where most of gene copy-number increases occurred on chromosome II (Borneman *et al.* 2011).

The genomic differences found in the natural isolates, identified both by SNP analysis and aCGH, may provide the basis for novel phenotypic characteristics. In order to further investigate this link, a phenotypic screen was devised to evaluate specific patterns for a set of physiological tests, including also tests that are important for winemaking strain selection. This experimental plan was previously applied with success for the characterization of several strains from different origins (Mendes *et al.* 2013), and was based on approaches that are generally applied for the selection of winemaking strains (Mannazzu, Clementi and Ciani 2002). Our results showed phenotypic differences in 14 tests from the 30 considered, being able to distinguished natural isolates from the reference strain (Table 3). In three tests, all the four natural isolates presented discriminatory results, which distinguished them from the reference strain: capacity to ferment must at 18°C, and inability to grow in the presence of $CuSO_4$ (5 mM) and SDS (0.01% w/v). Copper has been used for a long time as an antimicrobial agent in vineyards. Although copper resistance has been previously suggested as a consequence of environmental adaptation, arisen through positive selection, our results show that original VL1 strain had a slightly higher copper resistance compared to the reisolated strains. This seems somehow contradictory, since copper was used in the vineyards from where these strains were obtained. The resistance to the detergent SDS has been previously reported in wine strains (Kvitek, Will and Gasch 2008), which is in agreement with the use of detergents in the washing of fermentation vessels. This resistance was not shared by the natural isolates. Our findings are in agreement with previously reported generation of intrastrain phenotypic variability (Kvitek, Will and Gasch 2008; Camarasa *et al.* 2011; Mendes *et al.* 2013), that occur in altered environmental conditions, and that was associated with differences in the genomic expression patterns. In these studies, some phenotypes were able to distinguish groups of strains according to the ecological niches, providing evidence for phenotypic evolution driven by environmental adaptation to different conditions. For example, Kvitek, Will and Gasch (2008) compared gene copy-number variations and phenotypic profiles during stress resistance in *S. cerevisiae* strains, and described positive relations between genomic alterations and the degree of phenotypic alterations.

The observed phenotypic differences were also evident when the metabolomic profiles obtained at the end of must fermentations of VL1 isolates were compared, HPLC analysis revealed statistical significant differences regarding the production of malic acid, ethanol and/or glycerol among some natural isolates in comparison to the reference strain (Fig. 5). Isolate-dependent differences regarding aromatic profiles were obtained by GC-MS analysis (Fig. 6). The corresponding PCA showed that three alcohols differentiated the natural isolates from the reference strain: benzene ethanol (= 2-phenylethanol), 2-methyl-1-butanol and isobutanol (= 2-methyl-1-propanol), due to their presence in different concentrations at the end of the fermentation. These compounds are three of the major fusel alcohols produced during must fermentation, resulting from transamination of the corresponding amino acid in the Ehrlich pathway. In the present work, these alcohols were increased in the end of the fermentation performed by the commercial reference strain, being a differentiating factor among the natural isolates in which just one or two of these three compounds appeared to be increased. The VL1 reference strain, as a commercialized strain used in winemaking, should have the capacity to produce compounds with favorable aromatic contributions. Benzene ethanol and 2-methy-1-butanol are desired in finished wines due to their odor descriptors as roses, sweet, fragrant, flowery and honey-like for the first (Meilgaard 1975; Ferreira *et al.* 2000; Silva-Ferreira, Guedes de Pinho and De 2003; Cullere *et al.* 2004; Escudero *et al.* 2004; Siebert *et al.* 2005) and banana, sweet, aromatic and cheese in the case of the second one (Meilgaard 1975; Escudero *et al.* 2004; Moreno *et al.* 2005). On the contrary, 2-methyl-1-propanol is a non-desired alcohol in the end of the fermentation and has odor descriptors related to alcohol aroma, estery and fusel odors (Meilgaard 1975; Etiévant and Etievant 1991; Diedericks 1996). This compound revealed to be discriminating between the reference strain and the natural isolates mainly VL1–099 and VL1–020.

In conclusion, our results showed that isogenic isolates of the commercial wine yeast strain Zymaflore VL1 recovered from nature present genetic differences in comparison with the reference strain. We identified ORFs amplification, with an apparent stochastic distribution, corresponding to *Ty* elements and also to gene amplifications with various functions that could reflect adaptive mechanisms to environmental conditions. One of these amplified genes was *ASP3–2*, which is related with previous reports of increased expression during nitrogen starvation. Some SNPs were also identified in natural isolates and these differences could be related to mechanisms involved in the generation of intrastrain phenotypic variability, evidenced by dissimilarities identified in 14 phenotypic tests, and in the metabolomic profiles of must fermentations accomplished by VL1 isolates. These isolates beside some adaptation to the environmental conditions present already some diminished capacities related with winemaking, in comparison with the reference strain.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSYR online.

FUNDING

Ricardo Franco-Duarte and Inês Mendes are recipients of a fellowship from the Portuguese Science Foundation, FCT (SFRH/BD/48591/2008, SFRH/BD/74798/2010, respectively). Financial support was obtained from FEDER funds through the program COMPETE,by national funds through FCT by the projects FCOMP-01–0124–008775 (PTDC/AGR-ALI/103392/2008) and PTDC/AGR-ALI/121062/2010, and through the strategic funding UID/BIA/04050/2013.

Conflict of interest. None declared.

REFERENCES

Adams BG. Induction of galactokinase in *Saccharomyces cerevisiae*: kinetics of induction and glucose effects. *J Bacteriol* 1972;**111**:308–15.

Adams J, Puskas-Rozsa S, Simlar J, *et al.* Adaptation and major chromosomal changes in populations of *Saccharomyces cerevisiae*. *Curr Genet* 1992;**22**:13–9.

Akao T, Yashiro I, Hosoyama A, *et al.* Whole-genome sequencing of sake yeast *Saccharomyces cerevisiae* Kyokai no. 7. *DNA Res* 2011;**18**:423–34.

Argueso JJL, Carazzolle MF, Mieczkowski PA, *et al.* Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res* 2009;**19**:2258–70.

Bisson LF. Geographic origin and diversity of wine strains of *Saccharomyces*. *Am J Enol Vitic* 2012;**63**:165–76.

Bleykasten-Grosshans C, Friedrich A, Schacherer J. Genome-wide analysis of intraspecific transposon diversity in yeast. *BMC Genomics* 2013;**14**:399.

Bleykasten-Grosshans C, Neuvéglise C. Transposable elements in yeasts. *C R Biol* 2011;**334**:679–86.

Bon EPS, Carvajal E, Stanbrough M, *et al.* Asparaginase II of *Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* 1997;**63–65**:203–12.

Borneman AR, Desany BA, Riches D, *et al.* Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet* 2011;**7**:e1001287.

Borneman AR, Forgan AH, Pretorius IS, *et al.* Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain. *FEMS Yeast Res* 2008;**8**:1185–95.

Borneman AR, Pretorius IS, Chambers PJ. Comparative genomics: a revolutionary tool for wine yeast strain development. *Curr Opin Biotechnol* 2013;**24**:1–8.

Brion C, Ambroset C, Sanchez I, *et al.* Differential adaptation to multi-stressed conditions of wine fermentation revealed by variations in yeast regulatory networks. *BMC Genomics* 2013;**14**:681.

Camarasa C, Sanchez I, Brial P, *et al.* Phenotypic landscape of *Saccharomyces cerevisiae* during wine fermentation: evidence for origin-dependent metabolic traits. *PLoS One* 2011;**6**: e25147.

Carreto L, Eiriz MF, Gomes AC, *et al.* Comparative genomics of wild type yeast strains unveils important genome diversity. *BMC Genomics* 2008;**9**:524.

Carro D, Garcia-Martinez J, Pérez-Ortin JE, *et al.* Structural characterization of chromosome I size variants from a natural yeast strain. *Yeast* 2003;**20**:171–83.

Cullere L, Escudero A, Cacho J, *et al.* Gas chromatography-olfactometry and chemical quantitative study of the aroma of six premium quality spanish aged red wines. *J Agr Food Chem* 2004;**52**:1653–60.

Damon C, Vallon L, Zimmermann S, *et al.* A novel fungal family of oligopeptide transporters identified by functional metatranscriptomics of soil eukaryotes. *ISME J* 2011;**5**:1871–80.

Dequin S, Casaregola S. The genomes of fermentative *Saccharomyces*. *C R Biol* 2011;**334**:687–93.

Diedericks W. Static headspace analysis of beer volatiles. Tracking the three 'bad' flavors. *Peak* 1996;**3**:2–4.

Diezmann S, Dietrich FS. *Saccharomyces cerevisiae*: population divergence and resistance to oxidative stress in clinical, domesticated and wild isolates. *PLoS One* 2009;**4**:e5317.

Doniger SW, Kim HS, Swain D, *et al.* A catalog of neutral and deleterious polymorphism in yeast. *PLoS Genet* 2008;**4**:e1000183.

Dowell RD, Ryan O, Jansen A, *et al.* Genotype to phenotype: a complex problem. *Science* 2010;**328**:469.

Dunham MJ, Badrane H, Ferea T, *et al.* Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *P Natl Acad Sci USA* 2002;**99**:16144–9.

Dunn B, Levine RP, Sherlock G. Microarray karyotyping of commercial wine yeast strains reveals shared, as well as unique, genomic signatures. *BMC Genomics* 2005;**6**:53.

Dunn B, Richter C, Kvitek DJ, *et al.* Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res* 2012;**22**:908–24.

Engel SR, Cherry JM. The new modern era of yeast genomics: community sequencing and the resulting annotation of multiple *Saccharomyces cerevisiae* strains at the *Saccharomyces* Genome Database. *Database (Oxford)* 2013;**2013**:bat012.

Escudero A, Gogorza B, Melús MA, *et al.* Characterization of the aroma of a wine from maccabeo. Key role played by compounds with low odor activity values. *J Agr Food Chem* 2004;**52**:3516–24.

Etiévant PX, Etievant P. Wine. In: *Volatile Compounds in Food and Beverages*. New York, US: Marcel Dekker, 1991, 483–546.

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

Ferreira V, López R, Cacho JFJ, *et al.* Quantitative determination of the odorants of young red wines from different grape varieties. *J Sci Food Agr* 2000;**80**:1659–67.

Franco-Duarte R, Mendes I, Gomes AC, *et al.* Genotyping of *Saccharomyces cerevisiae* strains by interdelta sequence typing using automated microfluidics. *Electrophoresis* 2011;**32**:1447–55.

Galeote V, Novo M, Salema-Oom M, *et al.* FSY1, a horizontally transferred gene in the *Saccharomyces cerevisiae* EC1118 wine yeast strain, encodes a high-affinity fructose/H +symporter. *Microbiology* 2010;**156**:3754–61.

Gasch AP, Spellman PT, Kao CM, *et al.* Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 2000;**11**:4241–57.

Goto-Yamamoto N, Kitano K, Shiki K, *et al.* SSU1-R, a sulfite resistance gene of wine yeast, is an allele of SSU1 with a different upstream sequence. *J Ferment Bioeng* 1998;**86**:427–33.

Greig D, Leu J-YY. Natural history of budding yeast. *Curr Biol* 2009;**19**:R886–90.

Howell KSS, Bartowsky EJJ, Fleet GHH, *et al.* Microsatellite PCR profiling of *Saccharomyces cerevisiae* strains during wine fermentation. *Lett Appl Microbiol* 2004;**38**:315–20.

Johnston M. Feasting, fasting and fermenting. *Trends Genet* 1999;**9525**:29–33.

Jones GE, Mortimer RK. Biochemical properties of yeast L-asparaginase. *Biochem Genet* 1973;**9**:131–46.

Kvitek DJ, Will JL, Gasch AP. Variations in stress sensitivity and genomic expression in diverse *Saccharomyces cerevisiae* isolates. *PLoS Genet* 2008;**4**:31–5.

League GP, Slot JC, Rokas A. The ASP3 locus in *Saccharomyces cerevisiae* originated by horizontal gene transfer from *Wickerhamomyces*. *FEMS Yeast Res* 2012;**12**:859–63.

Legras J-L, Merdinoglu D, Cornuet J-M, *et al.* Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* 2007;**16**:2091–102.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;**25**:1754–60.

Liti G, Carter DM, Moses AM, *et al.* Population genomics of domestic and wild yeasts. *Nature* 2009;**458**:337–41.

Liti G, Schacherer J. The rise of yeast population genomics. *C R Biol* 2011;**334**:612–9.

Mannazzu I, Clementi F, Ciani M. Strategies and criteria for the isolation and selection of autochthonous starter. In: Ciani M (ed.). *Biodiversity and Biotechnology of Wine Yeasts*. Trivandrum: Research Signpost, 2002, 19–35.

Matsumoto K, Oshima Y. Isolation and characterization of dominant mutations resistant to carbon catabolite repression of galactokinase synthesis in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1981;**1**:83.

Meilgaard MC. Flavor chemistry of beer: Part II: flavor and threshold of 239 aroma volatiles. *MBAA Tech Q* 1975;**12**:151–68.

Mendes I, Franco-Duarte R, Umek L, *et al.* Computational models for prediction of yeast strain potential for winemaking from phenotypic profiles. Schacherer J, ed. *PLoS One* 2013;**8**:e66523.

Moreno JA, Zea L, Moyano L, *et al.* Aroma compounds as markers of the changes in sherry wines subjected to biological ageing. *Food Control* 2005;**16**:333–8.

Muller LAH, McCusker JHJ. Microsatellite analysis of genetic diversity among clinical and nonclinical *Saccharomyces cerevisiae* isolates suggests heterozygote advantage in clinical environments. *Mol Ecol* 2009;**18**:2779–86.

Nijkamp JF, van den Broek M, Datema E, *et al.* De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113–7D, a model for modern industrial biotechnology. *Microb Cell Fact* 2012;**11**:36.

Novo M, Bigey F, Beyne E, *et al.* Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *P Natl Acad Sci USA* 2009;**106**:16333–8.

Only U, Practices B, Analysis D, *et al.* *Illumina*. *Mate Pair Library v2 Sample Preparation Guide For 2–5 kb Libraries*, 2009.

Pérez-Ortin JE, Querol A, Puig S, *et al.* Molecular characterization of a chromosomal rearrangement involved in the adaptive evolution of yeast strains. *Genome Res* 2002;**12**:1533–9.

Rachidi N, Barre P, Blondin B. Multiple Ty-mediated chromosomal translocations lead to karyotype changes in a wine strain of *Saccharomyces cerevisiae*. *Mol Gen Genet* 1999;**261**:841–50.

Richards KD, Goddard MR, Gardner RC. A database of microsatellite genotypes for *Saccharomyces cerevisiae*. *Anton Leeuw* 2009;**96**:355–9.

Roberts IN, Oliver SG. The yin and yang of yeast: biodiversity research and systems biology as complementary forces driving innovation in biotechnology. *Biotechnol Lett* 2011;**33**:477–87.

Schacherer J, Ruderfer DM, Gresham D, *et al.* Genome-wide analysis of nucleotide-level variation in commonly used *Saccharomyces cerevisiae* strains. *PLoS One* 2007;**2**:e322.

Schacherer J, Shapiro JA, Ruderfer DM, *et al.* Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* 2009;**458**:342–5.

Schuller D, Alves H, Dequin S, *et al.* Ecological survey of *Saccharomyces cerevisiae* strains from vineyards in the vinho verde region of Portugal. *FEMS Microbiol Ecol* 2005;**51**:167–77.

Schuller D, Casal M. The genetic structure of fermentative vineyard-associated *Saccharomyces cerevisiae* populations revealed by microsatellite analysis. *Anton Leeuw* 2007;**91**:137–50.

Schuller D, Pereira L, Alves H, *et al*. Genetic characterization of commercial *Saccharomyces cerevisiae* isolates recovered from vineyard environments. *Yeast* 2007;**24**:625–36.

Schuller D, Valero E, Dequin S, *et al*. Survey of molecular methods for the typing of wine yeast strains. *FEMS Microbiol Lett* 2004;**231**:19–26.

Siebert TE, Smyth HE, Capone DL, *et al*. Stable isotope dilution analysis of wine fermentation products by HS-SPME-GC-MS. *Anal Bioanal Chem* 2005;**381**:937–47.

Silva-Ferreira A, Guedes de Pinho P, De U. Analytical method for determination of some aroma compounds on white wines by solid phase microextraction and gas chromatography. *J Food Sci* 2003;**68**:2817–20.

St John TP, Davis RW. The organization and transcription of the galactose gene cluster of Saccharomyces. *J Mol Biol* 1981;**152**:285–315.

Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *P Natl Acad Sci USA* 2001;**98**:5116–21.

Umek L, Zupan B, Schuller D, *et al*. Computational approaches for the genetic and phenotypic characterization of a *Saccharomyces cerevisiae* wine yeast collection. *Yeast* 2009;**26**: 675–92.

Valero E, Schuller D, Gambon B, *et al*. Dissemination and survival of commercial wine yeast in the vineyard: a large-scale, three-years study. *FEMS Yeast Res* 2005;**5**:959–69.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;**38**:e164.

Warringer J, Zörgö E, Cubillos FA, *et al*. Trait variation in yeast is defined by population history. *PLoS Genet* 2011;**7**:e1002111.

Wei W, McCusker JH, Hyman RW, *et al*. Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *P Natl Acad Sci USA* 2007;**104**:12825–30.

Wenger JW, Schwartz K, Sherlock G. Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*. *PLoS Genet* 2010;**6**:e1000942.

Winzeler EA, Castillo-davis CI, Oshiro G, *et al*. Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. *Genetics* 2003;**163**:79–89.

Yocum R (Jr), Hanley S, West R, *et al*. Use of lacZ fusions to delimit regulatory elements of the inducible divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1984;**4**: 1985.