Whole genome analysis of a wine yeast strain

Nicole C. Hauser, Kurt Fellenberg, Rosario Gil, Sonja Bastuck, Jörg D. Hoheisel and José E. Pérez-Ortí

Abstract

Saccharomyces cerevisiae strains frequently exhibit rather specific phenotypic features needed for adaptation to a special environment. Wine yeast strains are able to ferment musts, for example, while other industrial or laboratory strains fail to do so. The genetic differences that characterize wine yeast strains are poorly understood, however. As a first search of genetic differences between wine and laboratory strains, we performed DNA-array analyses on the typical wine yeast strain T73 and the standard laboratory background in S288c. Our analysis shows that even under normal conditions, logarithmic growth in YPD medium, the two strains have expression patterns that differ significantly in more than 40 genes. Subsequent studies indicated that these differences correlate with small changes in promoter regions or variations in gene copy number. Blotting copy numbers vs. transcript levels produced patterns, which were specific for the individual strains and could be used for a characterization of unknown samples.

Introduction

Saccharomyces cerevisiae wine yeast strains have been selected for more than 8000 years by human beings under conditions that favoured the evolution towards several specialized features, such as fast growth in high-sugar fruit juices, high yield of and resistance to ethanol and, more recently, sulphite resistance and good flavour production (reviewed in Querol and Ramón, 1996; Pretorius, 2000). Wine yeasts exhibit a greater variety in chromosome size and number than reported for laboratory strains. They are aneuploid (Kunkee and Bisson, 1993; Bakalinsky and Snow, 1990), with disomies or trisomies and, in some cases, they are near diploid or triploid. The aneuploidy can confer selective advantages by increasing the copy number of beneficial genes and protecting against lethal or deleterious mutations (Guijo et al., 1997; Bakalinsky and Snow, 1990). Aneuploidy and/or polyploidy are so widespread that the maintenance of an unbalanced chromosome set has been suggested to be advantageous (Sancho et al., 1986; Adams et al., 1992; Codón et al., 1998). Chromosomal changes include gain or loss of chromosomes, deletions (30–50 kb), presence of hybrid chromosomes (Bidenne et al., 1992; Rachidi et al., 1999) and, in most cases, duplications (30–390 kb) (Adams et al., 1992). It has been suggested that these rearrangements occurred by recombination through subtelomeric repeats and transposable elements (Ty elements) (Rachidi et al., 1999, and references therein). Minor differences, such as point mutations, are also possible. Such variations may also affect strain performance, if they localize inside an open reading frame (ORF) or regulatory regions of a gene, especially if it is a pathway regulating one (Cavaliere et al., 2000).

There are very few DNA sequences of wine yeasts published or stored in databases (Goto-Yamamoto et al., 1998; Masneuf et al., 1998) but it seems that,
overall, the sequence homology between the standard laboratory yeast strain S288c (whose genome was fully sequenced by Goffeau et al., 1996) and wine yeasts is higher than 99% (Masneuf et al., 1998; Pérez-Ortíz, unpublished observations). This fact should permit the systematic use in wine yeast of probes and primers obtained or designed from the S288c sequence (Puig et al., 1998, 2000a). DNA arrays made of S288c sequences have been used in recent years for the analysis of many cellular processes in yeast (DeRisi et al., 1997; Wodicka et al., 1997; Holstege et al., 1998; Jelinski and Samson, 1999), but usually with S288c or genetically related strains. However, when using long PCR products as probes for hybridization (DeRisi et al., 1997; Hauser et al., 1998; Cox et al., 1999), point mutations cannot be discriminated. Therefore, this kind of DNA array should also work in the analysis of non-laboratory yeast strains (Cavaleri et al., 2000).

In this study, the comprehensive gene arrays of Hauser et al. (1998) were used for the analysis of both the genomic copy number and transcript levels of wine yeast T73. Even under standard laboratory growth conditions, the transcription profiles of wine and laboratory strains differed significantly for several genes, some of which can be related to special features of wine yeast physiology. Three such genes, \( YHB1 \), \( SSU1 \) and \( YJL217w \), have been analysed in greater detail. Also, both transcript and genome analyses show that wine strain T73 is much less colonized by Ty transposons than S288c.

**Materials and methods**

**Strains and growth conditions**

The \( S. \ ceriseisiae \) yeast strains used in this study are FY1679 (\( MATa/z, \) \( ura3-52/ura3-52, \) \( trpl-\Delta1/\) \( TRP1, \) \( leu2-3/LEU2, \) \( his3-\Delta200/HIS3 \)) and MCY730 (\( MATa, \) \( ura3-52, \) \( lys2-801, \) \( ade2-101, \) \( trpl-\Delta1, \) \( his3-\Delta200 \)) as laboratory strains, as well as T73, a wine strain commercialized by Lallemand Inc. (Montreal, Canada) (Querol et al., 1992). Yeast cells were grown at 30°C in YPD rich medium (1% Yeast Extract, 2% Bacto-peptone, 2% glucose). Solid medium was supplemented with 2% agar.

**Nomenclature note**

We call ‘probes’ the pieces of DNA tethered to the array and ‘target’ the free, labelled DNA-fragments hybridized to the probes. Each probe-DNA was spotted in duplicate; the two spots are referred to as ‘primary’ and ‘secondary’ spot to indicate that they represent the same ORF.

**Nucleic acids isolation and hybridization**

Genomic DNA was isolated from yeast cells by phenol extraction with glass beads, essentially as described by Hoffman and Winston (1987) or with QIAGEN genomic DNA spin columns, as recommended by the manufacturer.

Total RNA from yeast cells was prepared as described by Sherman et al. (1986), but using a multiple-sample automated device (Fast-Prep, BIO101, Inc.) to break the cells, or following the procedure of Hauser et al. (1998).

All arrays were subjected to a standard control procedure prior to being used in experimentation. First, an oligonucleotide was hybridized that specifically bound to the forward primers of all PCR products. After removal of the oligonucleotide, total genomic DNA from a specific preparation was applied as a second control. While the former hybridization determined accurately the amount of DNA present at each spot, the latter experiment served as an indicator for the usefulness of the individual arrays in complex hybridizations.

Labelling of total RNA by reverse transcription was done as described (Hauser et al., 1998). Genomic DNA was sonicated to fragments and labelled by random priming. Hybridization with either target was performed according to the protocol used for labelled cDNA (Hauser et al., 1998).

**Signal acquisition and analysis**

Digital images of radioactive signals were acquired with Molecular Dynamics Storm 860 or Fuji FLA3000 phosphorimagers and quantified using the ArrayVision module of AIS (Imaging Research Inc.). Signal intensities of repeated hybridizations were normalized and significance levels assessed by two stringency criteria, as described by Beissbarth et al. (2000). The highly stringent ‘min–max separation’ is calculated by taking the minimum distance between all data points of the two strains. The less stringent criteria, called ‘standard deviation separation’, is defined as the difference of the means of the two data sets diminished by one standard deviation. In the tables, a colour code indicates the two stringency measures.
Northern blotting
Northern blot analysis was made on RNA samples that were separated by electrophoresis in formaldehyde-containing agarose gels (Sambrook et al., 1989). The DNA-transfer to nylon membranes (Hybond-N+, Amersham) was carried out according to the manufacturer’s instructions.

PCR amplification of DNA from strain T73 was made as follows: the PCR mixture was prepared with 200 ng genomic DNA as template, appropriate primer molecules (60 pmol each), 0.2 mM each dNTP, BioTaq buffer, 2.5 mM MgCl\textsubscript{2} and 2 U BioTaq (BioLine, UK) in a 50 \( \mu \)L volume. The reaction conditions were: 25 cycles of 30 s at 94°C, 30 s at 50°C and 30 s at 72°C, followed by a final incubation of 3 min at 72°C. PCR products were isolated after agarose gel electrophoresis by the Gene Clean (Bio101 Inc.) procedure and used for probe labelling by random priming or DNA sequencing.

Results

Use of S288c arrays for wine yeast analysis
Several highly stringent hybridizations on different arrays were performed consecutively with random-primed genomic DNA from laboratory and wine strains. The total number of detectable spots was identical within experimental variation, documenting that no major differences exist between the T73 wine yeast and the S288c genetic background. We concluded that overall the existing arrays are well suited for the analysis of yeast strains other than S288c. However, some reproducible differences in the intensity of several individual spots could be seen, as will be discussed below.

Comparative transcriptional analysis of yeast strains MCY730 and T73
Total RNA was obtained from either strain T73 or strain MCY730 (which has a S288c genetic background) grown with orbital agitation in YPD medium to middle logarithmic phase (OD\textsubscript{600} = 0.5 – 0.6). Repeated hybridizations were performed with material of both strains, resulting in six data sets each, considering primary and secondary spots individually. Relative changes between T73 and MCY730 and a measure of the significance level were calculated, the latter indicated by a colour code, as shown in Table 1. The complete list for all ORFs as well as the median of normalized signal intensities can be downloaded from our web pages: http://scsie.uv.es/chipsdna/ and http://www.dkfz-heidelberg.de/funct_genome/index.html. An overview of the transcript levels is presented in Figure 1. A list of selected genes, which either exhibited the most significant changes or proved to be of particular interest, is given in Table 1.

All of the Ty1 and Ty2 elements present on the array showed higher transcript levels in the laboratory strain MCY730 compared to T73 (Table 2). Every Ty encodes for two genes called TyA and TyB (Boeke and Sandmeyer, 1991). Since the Ty1 transposons are very similar in DNA sequence, the signal at each spot could be considered as an average of the contribution of all these transcripts. We observed much higher differences for Ty1B ORFs compared to Ty1A ORFs, the average ratio MCY730 : T73 being 38.8 and 8.6, respectively. Moreover, not only relative transcriptional changes are higher for Ty1B, but also the normalized signal intensities (see Table 2). The other Ty elements, except for one case of Ty2B, gave rise to signal intensities below the significance threshold and thus were not considered.

Another gene that appears to be higher expressed in MCY730 is URA3. Since the strain is ura\textsuperscript{−}, it is not able to synthesize uracil. However, the mutation ura3-52 is a transposon insertion into the ORF of the wild-type gene (Brachmann et al., 1998) and does not block transcription. This explains why a signal was detected. URA1, which functions downstream in the biosynthesis of uracil, also produced higher transcript levels in the ura3\textsuperscript{−} laboratory strain. Its transcription could be affected by a regulation deficiency in the biochemical pathway because of the lack of one enzyme. Similar explanations could be true for the genes involved in lysine and adenine biosynthesis, because the MCY730 strain is lys2\textsuperscript{−} and ade2\textsuperscript{−}. Similar results have been described for the BY4743 strain, which is his3\textsuperscript{−} and leu2\textsuperscript{−} (Hughes et al., 2000).

Two genes coding for a-factor specific proteins, MFa1 and MFa2, also exhibited high expression in MCY730. They are a-specific and should therefore be expressed in MATa strains, such as MCY730, and not expressed in diploid strains, such as T73. It is interesting to note that, in spite of comparing a haploid strain (MCY730) to a near-diploid one (T73), we did not find significant differences.
between the genes described by Galitski et al. (1999) as differentially expressed in dependence on the level of ploidy.

Another 23 of the genes shown in Table 1 are known to be involved in a variety of functions such as amino acid biosynthesis, purine biosynthesis and stress responses. Two of them seemed to be particularly interesting, SSU1 and YHB1. SSU1 encodes for a plasma membrane protein involved in sulphite resistance (Park and Bakalinsky, 2000). This gene had been previously shown to be differentially expressed in wine yeast strains other

### Table 1. Genes showing most significant transcriptional changes in wine yeast T73

<table>
<thead>
<tr>
<th>ORF ID</th>
<th>Gene</th>
<th>Predicted function</th>
<th>Rel changes T73 vs. MCY730</th>
<th>Signal intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAR015W</td>
<td>ADE1</td>
<td>Phosphoribosylamidimidazole-succinocarboxamide synthase</td>
<td>-4.3</td>
<td>4548.86</td>
</tr>
<tr>
<td>YAR068W</td>
<td></td>
<td></td>
<td>+3</td>
<td>1870.86</td>
</tr>
<tr>
<td>YBL113C</td>
<td></td>
<td>Subtelomeric ORF</td>
<td>-2.9</td>
<td>3697.81</td>
</tr>
<tr>
<td>YBR248C</td>
<td>HIS7</td>
<td>Glutamine amidotransferase/cyclase</td>
<td>-2.1</td>
<td>1555.90</td>
</tr>
<tr>
<td>YDL131W</td>
<td>LYS21</td>
<td>Homocitratre synthase</td>
<td>-2.7</td>
<td>8683.12</td>
</tr>
<tr>
<td>YDL182W</td>
<td>LYS20</td>
<td>Homocitratre synthase</td>
<td>-3.8</td>
<td>9816.50</td>
</tr>
<tr>
<td>YDR019C</td>
<td>GCV1</td>
<td>Glycine decarboxylase, subunit T</td>
<td>-2.5</td>
<td>1447.18</td>
</tr>
<tr>
<td>YDR461W</td>
<td>MFA1</td>
<td>Mating pheromone α-factor 1</td>
<td>-4.8</td>
<td>5301.49</td>
</tr>
<tr>
<td>YDR513W</td>
<td>TTR1</td>
<td>Guatereoxin</td>
<td>+1.6</td>
<td>2523.61</td>
</tr>
<tr>
<td>YDR529C</td>
<td>QCR7</td>
<td>Ubiquinol-cytochrome-c reductase subunit 7</td>
<td>+1.7</td>
<td>6684.11</td>
</tr>
<tr>
<td>YEL021W</td>
<td>URA3</td>
<td>Orotidine-5'-phosphate decarboxylase</td>
<td>-3.3</td>
<td>2813.68</td>
</tr>
<tr>
<td>YER034W</td>
<td></td>
<td></td>
<td>-9.4</td>
<td>9980.29</td>
</tr>
<tr>
<td>YFL068W</td>
<td></td>
<td></td>
<td>-2.4</td>
<td>7875.51</td>
</tr>
<tr>
<td>YGL255W</td>
<td>ZRT1</td>
<td>High-affinity zinc transport protein</td>
<td>+3.5</td>
<td>930.60</td>
</tr>
<tr>
<td>YGR234W</td>
<td>YHI1</td>
<td>Flavohaemoglobin</td>
<td>-10.1</td>
<td>1826.74</td>
</tr>
<tr>
<td>YHR018C</td>
<td>ARG4</td>
<td>Argininosuccinate lyase</td>
<td>-3.6</td>
<td>7829.88</td>
</tr>
<tr>
<td>YHR053C</td>
<td>CUP1A</td>
<td>Metallothionein</td>
<td>-7</td>
<td>36367.39</td>
</tr>
<tr>
<td>YHR055C</td>
<td>CUP1B</td>
<td>Metallothionein</td>
<td>-7.8</td>
<td>39818.48</td>
</tr>
<tr>
<td>YHR215W</td>
<td>PDO12</td>
<td>Secreted acid phosphatase</td>
<td>-4.4</td>
<td>4740.23</td>
</tr>
<tr>
<td>YIL051C</td>
<td>MMD1</td>
<td>Required for maintenance of mitochondrial DNA</td>
<td>-2.9</td>
<td>21835.35</td>
</tr>
<tr>
<td>YIL157C</td>
<td>FAR1</td>
<td>Cyclin-dependent kinase inhibitor (CKI)</td>
<td>-6.1</td>
<td>29549.83</td>
</tr>
<tr>
<td>YIL200C</td>
<td></td>
<td>Strong similarity to aconitase hydrolase</td>
<td>-3.1</td>
<td>3843.52</td>
</tr>
<tr>
<td>YIL217W</td>
<td></td>
<td></td>
<td>-5.4</td>
<td>14572.10</td>
</tr>
<tr>
<td>YPR025C</td>
<td>BNA1</td>
<td>3-Hydroxyanthranilic acid dioxygenase</td>
<td>-5.9</td>
<td>67579.81</td>
</tr>
<tr>
<td>YKL030W</td>
<td></td>
<td></td>
<td>+1.7</td>
<td>13559.74</td>
</tr>
<tr>
<td>YKL156W</td>
<td>RPS27A</td>
<td>Ribosomal protein S27.e</td>
<td>+1.7</td>
<td>7951.90</td>
</tr>
<tr>
<td>YKL161W</td>
<td>URA1</td>
<td>Dihydroorotate dehydrogenase</td>
<td>-3.7</td>
<td>7684.72</td>
</tr>
<tr>
<td>YKR080W</td>
<td>MTD1</td>
<td>Methyleneetetrahydrofolate dehydrogenase (NAD +)</td>
<td>+5.1</td>
<td>3345.43</td>
</tr>
<tr>
<td>YLR076C</td>
<td></td>
<td></td>
<td>+2.6</td>
<td>1513.76</td>
</tr>
<tr>
<td>YLR303W</td>
<td>MET25</td>
<td>0-Acetylaminoacridine sodium hydrolase</td>
<td>-1.7</td>
<td>18881.59</td>
</tr>
<tr>
<td>YLR339W</td>
<td>ADE13</td>
<td>Adenylosuccinate lyase</td>
<td>+6.1</td>
<td>6029.88</td>
</tr>
<tr>
<td>YMR173W</td>
<td>DQR4B</td>
<td>Heat shock protein</td>
<td>-6.9</td>
<td>7572.92</td>
</tr>
<tr>
<td>YMR236W</td>
<td>TAF17</td>
<td>TFIID and SAGA subunit</td>
<td>-2.4</td>
<td>9708.86</td>
</tr>
<tr>
<td>YNL055C</td>
<td>POR1</td>
<td>Mitochondrial outer membrane porin</td>
<td>-10.3</td>
<td>27585.02</td>
</tr>
<tr>
<td>YNL145W</td>
<td>MFA2</td>
<td>Mating pheromone α-factor 2</td>
<td>-11.4</td>
<td>10438.20</td>
</tr>
<tr>
<td>YNL306W</td>
<td>MRPS18</td>
<td>Ribosomal protein of the small subunit, mitochondrial</td>
<td>-2.1</td>
<td>4481.84</td>
</tr>
<tr>
<td>YNR050C</td>
<td>LYS9</td>
<td>Saccharopine dehydrogenase</td>
<td>-10.7</td>
<td>42463.49</td>
</tr>
<tr>
<td>YOL066W</td>
<td></td>
<td></td>
<td>-2.9</td>
<td>9345.35</td>
</tr>
<tr>
<td>YOL126C</td>
<td>MDH2</td>
<td>Malate dehydrogenase, cytoplasmic</td>
<td>-7</td>
<td>3063.58</td>
</tr>
<tr>
<td>YOL143C</td>
<td>ribbon</td>
<td>6,7-Dimethyl-8-ribitylumamine synthase</td>
<td>-9.8</td>
<td>27917.32</td>
</tr>
<tr>
<td>YOL092W</td>
<td>SSU1</td>
<td>Sulphite transport protein</td>
<td>+2.1</td>
<td>3074.89</td>
</tr>
<tr>
<td>YPR033W</td>
<td></td>
<td>Subtelomeric ORF</td>
<td>-3.2</td>
<td>3797.44</td>
</tr>
</tbody>
</table>

Factors of relative changes are highlighted with their respective colour of assessed quality measure: ● high-quality upregulation, ○ mid-quality upregulation, ● high-quality downregulation, ○ mid-quality downregulation.
than T73 (Goto-Yamamoto et al., 1998). YHBI1 encodes a flavohaemoglobin, whose expression is related to the presence of O2 (Liu et al., 2000). Both genes were analysed in greater detail (see below).

The remaining 10 differentially transcribed genes fall into the category of orphans. Two of them, YPR203w and YBL113c, belong to families of subtelomeric proteins, many members of which show partial homologies. Their signal intensities could, at least in part, be the result of cross-hybridization events between the transcripts of the members of their families. Another orphan gene, YJL217w, is analysed in greater detail below.

Figure 1. Comparison of transcript levels in strains T73 and MCY730. (A) Scatterplot of the median of the normalized signal intensities. Each black dot represents a gene. Indicated genes are presented as Northern blot analyses in Figure 2. (B) Presentation of the relative changes in transcript signals (top) and the normalized signal intensities (bottom) of particular ORFs. Dotted lines represent individual measurements, bold lines the respective mean value.
Northern hybridization analysis of selected genes

In order to confirm the results obtained with the arrays, we performed Northern hybridization analysis with probes specific for some of the selected genes. Figure 2 shows typical results for three genes: SSU1, YHB1 and YJL217w, with ACT1 acting as control. In all three cases, differences were observed, which are in agreement with the array results: SSU1 is higher expressed in T73; YHB1 and YJL217w were clearly expressed in MCY730 while their transcript levels were not detectable in T73.

Sequence analysis of some wine yeast promoters

There are several possible explanations for a gene to be differentially expressed in two strains. An obvious one is the presence of differences in the promoter sequence. To test this possibility, we chose few of the genes reported above. Oligonucleotide primers were designed to amplify by PCR the promoter regions from T73. The resulting fragments were then sequenced. In the case of YHB1, the sequence of the T73 promoter was very close to that of S288c, with only five differences in the amplified 577 bp region. All those differences were point mutations or single-base deletions (see Figure 3) except for one, an 8 bp deletion inside a 22 bp AT repeat placed 100 bp upstream of the ATG codon. Although there is no indication of a TATA box or other regulatory elements, it is tempting to speculate that the shortening of the AT repeat could explain the lower level of expression of YHB1 in T73. We also sequenced the SSU1 promoter. It exhibited a gross rearrangement (J. E. Pérez-Ortí and S. Puig, in preparation) identical to that observed for another wine strain, which was shown to increase sulphite resistance four-fold (Goto-Yamamoto et al., 1998).
Strain characterization by comparisons of genome vs. transcriptome

We plotted the values obtained with RNA targets (transcriptome) against genomic signal intensities (genome) (Figure 4) using data from successive hybridizations, using the very same filter array in order to avoid experimental bias as much as possible. In a scatterplot, such analysis produces a croissant-shaped cloud of dots that seems to be characteristic for each strain. Dots on the top-left corner of the cloud correspond to genes that are highly expressed but have low copy numbers (1–2 copies per haploid genome), such as ribosomal proteins or glycolytic enzymes. Dots on the bottom-right corner of the cloud correspond to genes with high values in genomic hybridization and low levels of expression; in Figure 4, we highlighted in this area a gene family composed of about 30 members with more than 90% similarity in protein sequence; YPR203w belongs to this family. Dots located at the top-right corner correspond to genes with both high copy number and high expression levels. Between the two strains analysed here, a clear pattern difference can be seen for the Ty1 family, for example most dots are clustered in top-right corner for FY1679 as opposed to those of T73, corroborating the results obtained for the MCY730/T73 comparison (Table 2).

Discussion

The initial goal of this study was to determine whether the yeast arrays generated with S288c DNA sequences (Hauser et al., 1998) were useful for the analysis of industrial yeast strains. For this purpose, the natural isolate T73 wine yeast strain (Querol et al., 1992) was used, with genomic DNA of the laboratory strains FY1679 and MCY730 acting as controls. In all cases, not all the probes produced a signal. However, the number of those that could be detected was very similar. From this, we conclude that most of the S288c genes are present in the wine strain T73. A similar conclusion can be obtained from a recent experiment on glass microarrays (Cavalieri et al., 2000). As a matter of course, it is possible that this particular wine strain or other industrial yeast strains contain genes which are not present at all or in a different copy number in the standard laboratory strains. This is known, for example, for most members of the families for SUC (Carlson and Botstein, 1983; Naumov et al., 1996), MEL (Naumov et al., 1990) and MAL (Naumov et al., 1994) and a gene that encodes for resistance to toxicity in molasses (Ness and Aigle, 1995). Most of these are translocated repeats of the original locus. However, they should represent only a very small fraction of the entire yeast genome. In our hands, a genomic hybridization of random-primed total DNA-samples proved to be a good procedure to detect differences in the quality of DNA arrays. This way of array validation is more sensitive than the hybridization of oligonucleotide tags, which bind to all the PCR products (Hauser et al., 1998).

The comprehensive gene arrays were used for the analysis of both the genomic copy number and transcript levels of wine yeast T73. RNA was obtained from yeast cells growing in middle logarithmic phase. This comparison has some possible drawbacks. The first is the growth rate. As usual for most industrial strains, T73 grows twice as fast as MCY730 (J. Gimeno, personal communication). This means that the selected conditions might not really be identical for both strains. However, this discrepancy is unavoidable when comparing strains that are so different. One
way to overcome this problem would be to perform the experiments using chemostat cultures, in which the growth rate of each strain could be controlled.

Another problem is the fact that MCY730 is a haploid strain while T73 is near-diploid (Puig et al., 1998, 2000b). This has been described to affect the expression of about a dozen genes (Galitski et al., 1999). We could not detect significant differences in this set of genes in our experiments. However, in our study both strains are of very different genetic background, rather than one being a duplication of the other, as was the case in the published study. This might have obscured ploidy differences.

Finally, some genes are known to be defective in MCY730. For instance, S288c background strains lack the transcription factor Flo8p (Liu et al., 1996). Whether or not this gene is active in T73 is unknown, but the flocculation properties of wine yeasts are very variable (discussed in Pretorius, 2000), so it is difficult to anticipate the differences.

Figure 3. Comparison of the promoter sequence of gene YH81 in T73 and the S288c sequence stored in the MIPS database (http://www.mips.biochem.mpg.de/). Sequence differences and the ATG codon are marked. The GenBank Accession No. for the YH81 T73 sequence is AF239759.
between MCY730 and T73. Our analysis does not show any significant variation in either \textit{FLO8} or its targets, except for the case of \textit{FLO1}, which is transcribed very poorly in T73 (32 times less than in MCY730), although the statistical confidence for this result only meets the less stringent quality criteria (see web pages). Other genes that differ between both strains are the ones related with \textit{MATa} mating type or with the MCY730 auxotrophies, such as \textit{ura3}⁻. These differences (see Results) serve as an internal control of data quality. For some of the selected genes, the results from the array analysis were confirmed by Northern blot experiments. In such cases, independently isolated RNA preparations were used in order to perform biologically meaningful control experiments. As described by others (Ter Linde \textit{et al.}, 1999), the overall expression differences found in array hybridizations were lower than those on Northern blots. In two cases (\textit{YHB1}, \textit{YJL217w}), we could not detect any signal on the T73 RNA samples.

We chose some particular genes for further analysis because of their special interest to wine yeasts. \textit{SSU1} has been demonstrated to code for a plasma membrane protein required for sulphite efflux (Park and Bakalinsky, 2000). Sulphite is a widely used preservative in wine production (Pretorius, 2000). The expression level of this gene directly correlates with sulphite resistance of the yeast strain (Goto-Yamamoto \textit{et al.}, 1998; Park and Bakalinsky, 2000). \textit{YHB1} codes for a flavohaemoglobin whose function is not well understood. It seems to protect cells against the damage caused by nitrosylation (Liu \textit{et al.}, 2000). Its expression is elevated in aerobic conditions (Liu \textit{et al.}, 2000; Zhao \textit{et al.}, 1996). Thus, this gene seems to be more important for cells growing in O₂-rich media. Its low expression in wine yeast may be relevant to physiological features of a strain, which has evolved for millions of generations under the O₂-limiting conditions of wine fermentation.

The main difference between the strains in Ty1 and Ty2 transposable elements seems to be the fact that the S288c laboratory genetic background has many more copies of Ty1 transposons than the T73 wine strain. This result confirms previous reports that Ty1 and Ty2 transposable elements are less frequent in wine and brewer’s yeasts than in laboratory and, especially, bakers’ strains (Codón \textit{et al.}, 1998). The observation is also compatible with the suggestion that in most or even all the Ty elements transposed recently, and in some situations perhaps in wine fermentations, a selective pressure against accumulation of Ty elements might exist (Jordan and McDonald, 1999). Since the differences at transcript level are more pronounced than at genomic level, it is possible that Ty transcription levels are either higher in each copy of Ty1 or that some Ty1 copies are silent in the wine strain. Ploidy may also significantly alter Ty

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Strain characterization by blotting genomic copy numbers vs. transcript levels. Signal intensities obtained from random primed genomic DNA are plotted vs. results of hybridizations with cDNA representing total RNA preparations. Each black dot represents one gene. Gene groups of particular interest are colour-coded as indicated.}
\end{figure}
expression, which is regulated by the mating pathway (reviewed in Boeke and Sandmeyer, 1991). On the other hand, the much higher labelling levels of Ty1B ORFs compared to Ty1A might reflect the fact that cDNA was synthesized from the 3' end of the mRNA. In Ty transposons, all the regular transcripts end just downstream of or inside the B-ORF (Boeke and Sandmeyer, 1991). Therefore, it is possible that more label was incorporated into this part of the sequence, since many of the retrotranscript molecules did not reach the 5' end of the mRNA. The fact that genomic samples did not show such an effect (Table 2) argues in favour of this hypothesis. However, this would only explain the results for MCY730 but not the ones obtained for T73, in which TyB labelling is not significantly different from TyA. In combination, all these results rather suggest that Ty transcription is in some way defective in T73-producing incomplete molecules. This would result in lower mRNA levels, especially for the 3'-portion of the message (TyB).

The comparison of genomic copy numbers vs. transcript levels (Figure 4) may be a general way to describe a given yeast strain. Especially the information on Ty elements and large gene families, such as the subtelomeric one shown in Figure 4, which are very variable between strains, could be used as a sensitive molecular tool for industrial strain identification, as has been suggested previously (discussed in Pretorius, 2000).

Acknowledgements

We thank Dr F. Estruch for the gift of strain MCY730 and Dr J. García-Martínez for his help with the analysis of gene families. We acknowledge the excellent sequencing work done by Sistemas Genómicos S.L. (Paterna, Spain) and the technical support by M. Bier. This work was funded by the European Commission as part of the EUROFAN-2 project under contract BIO4-CT97-2294 to J.E.P.-O. and J.D.H.; M.B. was supported by the Spanish programme, Ayudas para la Incorporación a España de Doctores y Tecnólogos, of the Ministerio de Educación y Cultura; J.E.P.-O. was the recipient of a Marie Curie Fellowship of the European Commission.

References


Functional Genome Analysis, DKFZ Heidelberg. URL: http://www.dkfz-heidelberg.de/funct_genome/index.html.


Jelinsky SA, Samson LD. 1999. Global response of Sacchar-
Whole genome analysis of a wine yeast strain

79

Copyright © 2001 John Wiley & Sons, Ltd.


Laboratorio de Chips de DNA de la Universitat de València. URL: http://scsie.uv.es/chipsdna/.


Submit your manuscripts at
http://www.hindawi.com