Microbiology (2009), **155**, 4036–4049 **DOI** 10.1099/mic.0.030726-0

Functional analyses of *PAU* genes in *Saccharomyces cerevisiae*

Zongli Luo and Hennie J. J. van Vuuren

Wine Research Centre, Faculty of Land and Food Systems, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

Correspondence
Hennie J. J. van Vuuren
hjjvv@interchange.ubc.ca

Received 8 May 2009
Revised 27 July 2009
Accepted 15 September 2009

PAU genes constitute the largest gene family in *Saccharomyces cerevisiae*, with 24 members mostly located in the subtelomeric regions of chromosomes. Little information is available about *PAU* genes, other than expression data for some members. In this study, we systematically compared the sequences of all 24 members, examined the expression of *PAU3*, *PAU5*, *DAN2*, *PAU17* and *PAU20* in response to stresses, and investigated the stability of all Pau proteins. The chromosomal localization, synteny and sequence analyses revealed that *PAU* genes could have been amplified by segmental and retroposition duplication through mechanisms of chromosomal end translocation and Ty-associated recombination. The coding sequences diverged through nucleotide substitution and insertion/deletion of one to four codons, thus causing changes in amino acids, truncation or extension of Pau proteins. Pairwise comparison of non-coding regions revealed little homology in flanking sequences of some members. All 24 *PAU* promoters contain a TATA box, and 22 *PAU* promoters contain at least one copy of the anaerobic response element and the aerobic repression motif. Differential expression was observed among *PAU3*, *PAU5*, *PAU17*, *PAU20* and *DAN2* in response to stress, with *PAU5* having the highest capacity to be induced by anaerobic conditions, low temperature and wine fermentations. Furthermore, Pau proteins with 124 aa were less stable than those with 120 or 122 aa. Our results indicate that duplicated *PAU* genes have been evolving, and the individual Pau proteins might possess specific roles for the adaptation of *S. cerevisiae* to certain environmental stresses.

INTRODUCTION

PAU genes are the largest gene family in *Saccharomyces cerevisiae*. The first *PAU* gene, *PAU1*, was isolated by accident and placed into the *PAU* gene family by sequence analysis and chromo-blotting (Viswanathan *et al.*, 1994). With the availability of complete genome sequences of *S. cerevisiae* strain S288C, 23 *PAU* members were initially identified (Goffeau *et al.*, 1996). An additional member was later revealed by comparative genome sequence analysis between *S. cerevisiae* and other yeast species (Blandin *et al.*, 2000), which finally brought the total number of *PAU* genes to 24. Among the 24 *PAU* genes, 15 members, YAL068C, YBL108C-A, YDR542W, YGL261C, YGR294W, YHL046C, YIL176C, YIR041W, YKL224C, YLL025W, YLL064C, YMR325W, YOL161C, YOR394W and YPL282C, have not officially been annotated, and their names were only recently reserved in the Saccharomyces Genome Database (SGD) as *PAU8–PAU22*.

PAU genes usually encode almost identical proteins of 120–124 aa. Due to high sequence homology, functional analysis of such a complex family has been challenging. Nevertheless, the induction of the expression of some *PAU* genes by a range of environmental factors has been reported (Rachidi *et al.*, 2000; Davies & Rine, 2006; Abramova *et al.*, 2001b; Homma *et al.*, 2003; Sahara *et al.*, 2002). The observations that *PAU* genes are highly upregulated during early stages of wine fermentation (Rossignol *et al.*, 2003; Marks *et al.*, 2008) and that the intracellular concentration of Pau5p is regulated by wine-making stress (Luo & van Vuuren, 2008) suggest that *PAU* genes play roles in the adaptation of yeast cells to stress during alcoholic fermentations. Furthermore, Pau proteins are homologous to N-terminal sequences of Tir/Dan proteins, which have a C-terminal putative glycophospholipid (GPI) attachment signal. Tir/Dan proteins might play a role in maintaining cell wall integrity or sterol uptake during stress (Abramova *et al.*, 2001a; Alimardani *et al.*, 2004).

Most of the *PAU* genes are located in the subtelomeric region of chromosomes, a place for maintaining gene families that serve adaptive purposes (Fabre *et al.*, 2005). Subtelomeric regions have been evolving more quickly than the rest of the genome (Kellis *et al.*, 2003). The other large subtelomeric gene family in *S. cerevisiae* is the *DUP* gene family, which has 23 members. *DUP* gene duplication

Abbreviations: PGK, 3-phosphoglycerate kinase; SGD, Saccharomyces Genome Database; TAP, tandem affinity purification.

is thought to have occurred through non-reciprocal translocation, tandem duplication and transposon Ty-associated duplication (Despons *et al.*, 2006). Gene duplication has long been considered a major force driving evolution (Ohno, 1970), and may occur via three mechanisms: chromosomal or segmental duplication, retroposition and tandem duplication (Zhang, 2003; Kong *et al.*, 2007; Pan & Zhang, 2008). The duplicated genes may either be lost or gain novel functions to retain their existence through divergence of sequence and expression during evolution (Taylor & Raes, 2004; Lynch & Conery, 2000). Systematic analyses at the genome level have shown that a large proportion of duplicated genes in yeast have diverged rapidly, and the vast majority of gene pairs eventually become divergent in sequence and expression (Gu *et al.*, 2002). How so many *PAU* genes originated, why they were preserved and evolved, and their specific roles, are still unknown.

With these questions in mind, we systematically compared the regulatory and coding sequences of all 24 members, analysed induction of *PAU3*, *PAU5*, *DAN2*, *PAU17* and *PAU20* in response to a range of environmental factors, and characterized the stability of 17 Pau proteins. We observed that *PAU* genes might have originated from *TIR/DAN* genes, been amplified through Ty-associated recombination and chromosomal end translocation, and diverged in sequence, stress response and protein stability.

METHODS

Sequence analyses. Sequences were derived from the SGD (http://www.yeastgenome.org). The alignment method was CLUSTAL W in T-COFFEE (Notredame *et al.*, 2000) v.5.05 or MEGA 3.1 (Kumar *et al.*, 2004). The program was run using default options. The output was viewed with Jalview and inspected visually. A phylogenetic tree (neighbour-joining method) was established with MEGA 3.1. The similarity of sequences was examined using BLASTN2.2.21+ (Altschul *et al.*, 1997). The YEASTRACT database (Teixeira *et al.*, 2006) was searched for DNA motifs in promoter regions.

Yeast strains and culture conditions. *S. cerevisiae* strain BY4742 (*MATα his3Δ1*, *leu2Δ0*, *lys2Δ0*, *ura3Δ0*) was used in this study to amplify the *PAU* genes, and to construct strains that encode tandem affinity purification (TAP) sequences or TAP-tagged Pau proteins in the genome (Table 1) or on plasmids. A wild-type diploid wine yeast strain S92 was also used for amplification of the *PAU7*, *PAU9* and *PAU21/22* genes, and for expression of TAP-tagged genes and wine fermentation. All yeast strains were grown in Difco yeast extract-peptone-dextrose (YPD; 1 % yeast extract, 2 % peptone, 2 % glucose) or minimal medium [0.67 % Difco yeast nitrogen base (YNB), 2 % glucose plus the required amino acids]. For stress induction, cells were grown aerobically in YPD at 25 °C to mid-exponential phase, centrifuged, washed with sterilized water, and then resuspended in YPD, YNB (without a carbohydrate, amino acids and ammonium sulphate for combination of nitrogen and carbohydrate starvation stress), YNB plus 2 % glucose (for nitrogen starvation stress), YNB plus amino acid and ammonium sulphate (for carbohydrate starvation stress), or YNB plus 2 % galactose. Cells were then grown aerobically at 25 °C for 2 h or 10 °C for 24 h. To study Pau–TAP induction during anaerobic growth, cells were grown in 5 ml YPD in a tightly capped 5 ml tube without shaking at 25 °C. Anaerobic conditions for testing the activity of the *PAU* promoters were obtained using BBL GasPak anaerobic systems (Becton Dickinson).

Table 1. Yeast strains used in this study

Strain	Genotype	Source
BY4742	*MATα his3Δ1*, *leu2Δ0*, *lys2Δ0*, *ura3Δ0*	Invitrogen
LY5	BY4742, *PAU3-TAP-URA3*	This study
LY6	BY4742, *PAU5-TAP-URA3*	This study
LY7	BY4742, *PAU17-TAP-URA3*	This study
LY8	BY4742, *PAU20-TAP-URA3*	This study
LY9	BY4742, *DAN2-TAP-URA3*	This study
LY19	BY4742, *pau5*::*TAP-KanMX4*	This study
LY20	BY4742, *pau3*::*TAP-KanMX4*	This study
LY280	BY4742, *pau17*::*TAP-URA3*	This study
LY281	BY4742, *pau20*::*TAP-URA3*	This study
LY282	BY4742, *dan2*::*TAP-URA3*	This study
S92	*MATα/a*	BioSpringer USA
LY12-7	S92, *PAU3-TAP-KanMX4*	This study
LY15-8	S92, *PAU5-TAP-KanMX4*	This study
LY21-7	S92, *DAN2-TAP-KanMX4*	This study
LY22-8	S92, *PAU17-TAP-KanMX4*	This study
LY23-8	S92, *PAU20-TAP-KanMX4*	This study

TAP tagging and plasmids. Genome TAP tagging was performed as described previously (Luo & van Vuuren, 2008), and the oligonucleotides used for tagging are listed in Table 2. The following plasmids were constructed: (1) pUG36-*PAU5*. The coding region of the *PAU5* gene was amplified with primers p89 and p148 from genomic DNA of strain BY4742, digested with *Spe*I and *Xho*I, and then cloned into *Spe*I–*Xho*I sites of pUG36. (2) pUG36-*PAU7* (B) and pUG36-*PAU7* (S). Due to the high degree of sequence homology with some of the other *PAU* genes, the coding sequence of *PAU7* cannot be amplified directly by one round of PCR. We therefore first amplified the specific sequences covering the *PAU7* ORF by using primers p89 and p141 from laboratory strain BY4742 and wine strain S92, respectively. The *PAU7* coding sequence was then amplified from the first-round PCR products with primers p89 and p151. After digestion with *Spe*I and *Xho*I, the PCR products were cloned into *Spe*I–*Xho*I sites of pUG36, yielding pUG36-*PAU7* (B) (B refers to the *PAU7* coding sequence from strain BY4742) and pUG36-*PAU7* (S) (S refers to the *PAU7* coding sequence from strain S92). (3) pUG36-*PAU9*. For the same reason as above, a two-round PCR strategy was used to clone the *PAU9* ORF. The first round was carried out using primers p142 and p143 from BY4742 and primers for the second-round PCR were p89 and p151. The PCR products were trimmed by *Spe*I and *Xho*I, and inserted into pUG36 pre-digested with the same enzymes. (4) p*PAU21/22*-TAP, p*PAU21/22* (*M1I*)-TAP and p*PAU21/22* (*M41I*)-TAP. The entire *PAU21/22* coding sequence plus their identical 1018 bp promoter sequence was amplified by primers p146 and p147 from genomic DNA extracted from BY4742. After digestion with *Xba*I and *Xho*I, the fragment was ligated into the *Spe*I and *Xho*I sites of pKS, yielding plasmid pKS-*PAU21/22*, on which the mutated *PAU21/22* gene encoding either the M1I or the M41I substitution was generated by site-directed mutagenesis using a QuickChange mutagenesis kit (Stratagene). After the mutation was confirmed by DNA sequencing, an *Xba*I–*Xho*I fragment of the wild-type and the mutated *PAU21/22* gene was inserted into the same sites of pRS326, yielding pRS326-*PAU21/22*, pRS326-*PAU21/22* (*M1I*) and pRS326-*PAU21/22* (*M41I*). Subsequently the TAP tag plus *CYC1* termination sequences were cut from plasmid p*MET25p*-TAP (Luo & van Vuuren,

Table 2. Oligonucleotides used in this study

Number and name	Sequence (5′–3′)
p17. PAU3-up	cagtgctctatctaaggacggtatctacactgctattccaaaatccatggaaaagagaag
p18. PAU5-up	ctctagcgctctatccgcagacggtatctacactattgcaaattccatggaaaagagaag
p19. DAN2-up	cagcgctctatccaaggacggtatctacactgctgttccaaactccatggaaaagagaag
p20. PAU17-up	cagcgctctatctgcagacggtatctacactgctgtaccaaactccatggaaaagagaag
p21. PAU20-up	ctccaaggctctatccaaagatggtatctacactatcgcaaactccatggaaaagagaag
p22. PAU3-down	cagcactactcattaagaatatatgcaataatactggatagcatacgactcactataggg
p23. PAU5-down	ttcctgcctagtgtatttcgtcagggcgaaccttcggataggctacgactcactataggg
p24. DAN2-down	aaatcataattatttttatccgttcgtaaagcaatccagtgcttacgactcactataggg
p25. PAU17-down	cgaaaccaatttcatttattcataacattctaatactcgtaaatacgactcactataggg
p26. PAU20-down	atttgaaatttataatttgctataaatcttggttttcttggtctacgactcactataggg
p27. PAU3-f	actctatctccatctgacgaaagg
p28. PAU5-f	gtgaccagtaatagcctgtttggg
p29. DAN2-f	cagctttgattaactctttccgac
p30. PAU17-f	catgcctcgttcaggaacaagtgc
p31. PAU20-f	catgcctcgttcaggaacaagtgc
p37. TEV-r	caaattctcgctagcagtagttgg
p47. TAP_C_XhoI	attgtactcgagtcaggttgacttccccg
p89. PAU3_speI	ctgcagactagtatggtcaaattaacttcaatcg
p90. DAN2_speI	ctgcagactagtatggtcaaactaacttcaattg
p91. PAU2_speI	ctgcagactagtatggtcaaattaacttcaattg
p141. PAU7_r	tcgtaaagcagggactctag
p142. PAU9_f	tcaggaatgagtaacaaagatg
p143. PAU9_r	tcattcgtaaaggcatcactcc
p144. PAU21_f	gagaatatgtgaagtgctaaag
p145. PAU21_r	tcgttcgtaaaagcttctctc
p146. PAU21_XbaI	attgtatctagaaatggggcatttgacac
p147. PAU3_XhoI	attgtactcgagttttggaatagcagtgtagatacc
p148. PAU5_XhoI	attgtactcgagatttgcaatagtgtagatacc
p149. DAN2_XhoI	attgtactcgaggtttggaacagcagtgtagatacc
p150. PAU17_Xho1	attgtactcgaggtttggtacagcagtgtagatacc
p151. PAU20_Xho1	attgtactcgaggtttgcgatagtgtagatacc
p153. DAN3_SpeI	ctgcagactagtatggtcaaactaacttcaatcgc
p166. PAU21_m1f	cgaatcgtttaagtggtgatcaccaatgaaggaatagg
p167. PAU21_m1r	cctattccttcattggtgatcaccacttaaacgattcg
p168. PAU21_m2f	ctaacaaacaaaacaaatacaatcgtcaaattaacttcaatcgctgctgg
p169. PAU21_m2r	ccagcagcgattgaagttaatttgacgattgtatttgttttgtttgttag
p180. PAU2_f	ggtatatgcctccactcttctc
p181. PAU2_r	ctaagtttcatagcatttccttcatt
p182. PAU4_fc_XbaI	ctgcagtctagaatggtcaaattaacttcaatcgcc
p183. PAU4_rc_XhoI	attgtactcgaggtttgcgatagtgtagatacca
p184. PAU16_r	gttcgtaaaattggcgtttgtctctaa
p185. PAU16_rc_XhoI	attgtactcgagatttgcgatagtgtagatacc
p186. PAU10_mf	ccagtgctctatccaaggtcggtatctacactatcg
p187. PAU10_mr	cgatagtgtagataccgaccttggatagagcactgg
p188. PAU12_mf	ccaatactactctttccaagccgctcatccaactgaaacc
p189. PAU12_mr	ggtttcagttggatgagcggcttggaaagagtagtattgg
p190. PAU13_rc_XhoI	attgtactcgaggtttgtgatagtgtagataccgtcc
p191. PAU15_mf	gtttcaagcagctcatccaagtgagacctacccagttgag
p192. PAU15_mr	ctcaactgggtaggtctcacttggatgagctgcttgaaac

2008) by *Sal*I and *Kpn*I and ligated to the *Xho*I–*Kpn*I sites of pRS326-*PAU21/22*, pRS326-*PAU21/22* (*M1I*) and pRS326-*PAU21/22* (*M41I*), yielding p*PAU21/22*-TAP, p*PAU21/22* (*M1I*)-TAP and p*PAU21/22* (*M41I*)-TAP, respectively. (5) *Pau–TAP* expression plasmids. To generate plasmids containing the Pau–TAP fusion expression cassette under the control of the *MET25* promoter, the intermediate cloning plasmids were first constructed. The coding sequences of *PAU2*, *PAU3*, *PAU4*, *PAU5*, *PAU7*, *PAU9*, *PAU13*, *PAU16*, *PAU17*, *PAU19*, *PAU20*, *DAN2* and *DAN3* were amplified from genomic DNA extracted from BY4742 or existing plasmids, digested with *Spe*I and *Xho*I, and ligated to the same sites of the cloning vector pKS, yielding pKS-*PAU2*, pKS-*PAU3*, pKS-*PAU4*, pKS-*PAU5*, pKS-*PAU7*, pKS-*PAU13*, pKS-*PAU16*, pKS-*PAU17*, pKS-*PAU19*, pKS-*PAU20*, PKS-*DAN2* and pKS-*DAN3*, respectively. pKS-*PAU1* was generated by

ligation of the small *Eco*RV–*Nco*I fragment of pKS-*PAU20* with the large *Eco*RV–*Nco*I fragment of pKS-*PAU9*. pKS-*PAU6* was generated by ligation of the small *Xba*I–*Nco*I fragment of pKS-*PAU9* with the large *Xba*I–*Nco*I fragment of pKS-*PAU20*. Plasmids pKS-*PAU10*, pKS-*PAU15* and pKS-*PAU12* were derived by site-directed mutagenesis of plasmids pKS-*PAU9*, pKS-*PAU19* and pKS-*DAN3*, respectively, based on only one amino acid difference between these pairs of Pau proteins. After DNA sequencing, the correct *PAU* ORFs were cut from their intermediate cloning plasmids with *Xba*I and *Xho*I and inserted into vector p*MET25p*-TAP pre-digested with the same restriction enzymes, yielding the following plasmids which encode different TAP-fused Pau proteins driven by the *MET25* promoter: p*MET25p-PAU1*-TAP, p*MET25p-PAU2*-TAP, p*MET25p-PAU3*-TAP, p*MET25p-PAU4*-TAP, p*MET25p-PAU5*-TAP, p*MET25p-PAU6*-TAP, p*MET25p-PAU7*-TAP, p*MET25p-PAU9*-TAP, p*MET25p-PAU10*-TAP, p*MET25p-PAU12*-TAP, p*MET25p-PAU13*-TAP, p*MET25p-PAU15*-TAP, p*MET25p-PAU16*-TAP, p*MET25p-PAU17*-TAP, p*MET25p-PAU19*-TAP, p*MET25p-PAU20*-TAP, p*MET25p-DAN2*-TAP and p*MET25p-DAN3*-TAP. All DNA manipulations were performed by standard procedures as described by Ausubel *et al.* (1999), and oligonucleotide sequences used for cloning are shown in Table 2.

Protein extraction and Western blot analyses. Whole-cell extracts were prepared and subjected to SDS-PAGE and immunoblotting as described previously (Luo & van Vuuren, 2008). A monoclonal anti-GFP antibody (Roche) was used to detect GFP-fusion proteins. Peroxidase–anti-peroxidase soluble complex (PAP; Sigma) was used to detect TAP or TAP-tagged proteins. As a loading control, Pgk1p was also detected in the blots using monoclonal anti-3-phosphoglycerate kinase (anti-PGK) antibody (Molecular Probes).

Cycloheximide chase assay. For aerobic chasing, yeast strains transformed with p*MET25p-PAU1*-TAP and p*MET25p-PAU3*-TAP were grown aerobically at 25 °C to mid-exponential phase in YNB plus essential nutrients. After addition of cycloheximide to a final concentration of 100 μg ml^{-1}, the cell culture was chased for 10, 15, 30 and 40 min, respectively. For anaerobic chasing, cells were grown in a 50 ml bottle filled with 40 ml YNB plus essential nutrients and fitted with an S-shaped air lock. When the cell density reached an OD_{600} 0.5, cycloheximide at a final concentration of 100 μg ml^{-1} was injected through the rubber stopper to the culture. After addition of cycloheximide, the culture was kept under anaerobic conditions, and samples were withdrawn using syringes at 10, 15, 30 and 40 min. Equal volumes of cells (1 OD_{600}) were collected, added to 10 mM NaN_3 and incubated on ice for 10 min. Cell lysis and immunoblotting were performed as previously described (Luo & van Vuuren, 2008).

RESULTS AND DISCUSSION

Expansion of *PAU* genes in *S. cerevisiae*

There are 24 *PAU* genes in *S. cerevisiae* distributed across all 16 chromosomes; 19 copies are located in subtelomeric regions, and five copies are found in internal regions (Fig. 1). Using the synteny viewer in SGD, four *PAU* genes (*PAU2*, *PAU11*, *PAU12* and *DAN2*) were detected in *Saccharomyces bayanus*, five (*PAU2*, *PAU7*, *PAU9*, *PAU11* and *PAU17*) in *Saccharomyces mikatae* and five (*PAU2*, *PAU7*, *PAU9*, *PAU21* and *DAN3*) in *Saccharomyces paradoxus*, far fewer than the 24 copies in *S. cerevisiae*. We also searched for *PAU* orthologues in the genomes of ancient yeast species that existed before whole-genome duplication through the yeast gene order browser (YGOB) (Byrne & Wolfe, 2006), and only found *TIR/DAN* genes, with 5′ end sequences that are highly similar to *PAU* genes. This observation raised the possibility that *PAU* genes originated from the N-terminal sequence of *TIR/DAN* genes. Inspection of the gene order and the synteny around all *PAU* genes in *S. cerevisiae* revealed some clues about how *PAU* genes expanded. Four of the five internal *PAU* genes (*PAU2*, *PAU5*, *PAU17* and *DAN2*) are flanked by a transposable element (Ty1) at one end, suggesting that they were integrated into the region by recombination between transposable elements. The other internal *PAU* gene, *PAU7*, is located adjacent to a transposition hot-spot where one solo and five tandem repeats from the *DUP240* family are located (Wirth *et al.*, 2005). It is interesting that all subtelomeric paralogues of *PAU* genes are in the same orientation relative to the telomere: genes close to the left end of chromosomes are located on the Crick strand, and genes close to the right end of chromosomes are on the Watson strand (Fig. 1). This same orientation implies that they might have been duplicated by chromosome end translocation. The reason why so many *PAU* genes have evolved remains a mystery. It has been proposed that subtelomeric regions are preferred sites for efficient amplification of useful genes, the expression of which gives yeast cells advantages to adapt to environmental changes (Fabre *et al.*, 2005). For example, it has been proposed that the *SUC*, *MAL* and *MEL* families were amplified in subtelomeric regions for yeast to use specific carbon sources (Pryde *et al.*, 1997). Given that *S. cerevisiae* is the main natural species used to ferment grape must to produce wine, and that the expression of *PAU* genes is induced during wine fermentation, the expansion of *PAU* genes in the subtelomeric region of *S. cerevisiae* could contribute to its adaptation to wine fermentation stress.

Sequence divergence in the coding region of *PAU* genes

During its evolution, *S. cerevisiae* has evolved mechanisms, such as sequence specialization, to acquire novel functions for duplicated genes. Although the exact role of *PAU* genes remains unclear, it is worthwhile to detect the sequence variation amongst the *PAU* members. The sequence alignment of *PAU* genes and Pau proteins was therefore performed using MEGA 3.1 (Kumar *et al.*, 2004) (Figs 2 and 3); 100 % identity was shown within the following pairs: *PAU1/PAU14*, *PAU6/PAU18*, and *PAU21/PAU22*. These pairs were located within the duplicated segments at the chromosome ends, which share high sequence similarity (above 90 %). Insertion or deletion of one to four codons was found to be the reason for length polymorphism of Pau proteins whose lengths are 120, 122, 123 or 124 aa (Fig. 3). A single-nucleotide substitution or insertion was observed to cause the truncation of Pau7p (55 aa) and Pau9p (42 aa), and the extension of Pau21/22p (164 aa). These mutations could be systematic sequencing errors as reported for other genes in SGD. The most frequently occurring mutations are

ORF/NAME	ORF length		Chromosomal location (distance (kb) from the telomere end)
	nt (bp)	AA	
YAL068C/PAU8	363	120	Chr I (2.1), subtelomeric
YAR020C/PAU7	168	55	Chr I (53.1), internal
YBL108C-A/PAU9	129	42	Chr II (7.7), subtelomeric
YBR301W/DAN3	363	120	Chr II (4.1), subtelomeric
YCR104W/PAU3	375	124	Chr III (8.8), subtelomeric
YDR542W/PAU10	363	120	Chr IV (8.6), subtelomeric
YEL049W/PAU2	363	120	Chr V (63.7), internal
YFL020C/PAU5	369	122	Chr VI (99.5), internal
YGL261C/PAU11	363	120	Chr VII (6.6), subtelomeric
YGR294W/PAU12	363	120	Chr VII (10.6), subtelomeric
YHL046C/PAU13	363	120	Chr VIII (12.2), subtelomeric
YIL176C/PAU14	363	120	Chr IX (9.1), subtelomeric
YIR041W/PAU15	375	124	Chr IX (5.9), subtelomeric
YJL223C/PAU1	363	120	Chr X (9.1), subtelomeric
YKL224C/PAU16	372	123	Chr XI (2.1), subtelomeric
YLL025W/PAU17	375	124	Chr XII (94.7), internal
YLL064C/PAU18	363	120	Chr XII (13.4), subtelomeric
YLR037C/DAN2	375	124	Chr XII (855.1), internal
YLR461W/PAU4	363	120	Chr XII (15.2), subtelomeric
YMR325W/PAU19	375	124	Chr XIII (1.7), subtelomeric
YNR076W/PAU6	363	120	Chr XIV (2.4), subtelomeric
YOL161C/PAU20	363	120	Chr XV (11.9), subtelomeric
YOR394W/PAU21	495	164	Chr XV (8.5), subtelomeric
YPL282C/PAU22	495	164	Chr XVI (8.4), subtelomeric

Fig. 1. Chromosomal location of *PAU* genes in *S. cerevisiae* strain S288C. The number in parentheses indicates the distance (in kb) of the *PAU* gene from the telomere end. Subtelomeric regions were defined by a distance of less than 20 kb from each telomere end (Wyrick *et al.*, 1999). AA, amino acid; vertical lines indicate the relative location of *PAU* genes on the chromosomes; the filled circle represents the centromere.

nucleotide substitutions. Synonymous single-nucleotide changes outnumber non-synonymous changes. A strong preference for C/T transition at the third place of the triplet codon was also observed. For example, compared with the consensus sequence, *PAU5* showed 48 nucleotide variations, of which 28 (58.3 %) were C/T transitions that do not change the amino acid sequence. This preference might explain why *PAU* genes exhibit highly biased codon usage: among 122 codons of *PAU5*, 87 contain C or T in the third position. Thus, *PAU* genes seem to have evolved with selection pressure to maintain their protein sequences.

Revisiting the sequences of *PAU7*, *PAU9* and *PAU21/22*

Given that some errors have been reported in the SGD sequence data, we reasoned that the single-nucleotide mutations which lead to truncation of Pau7p and Pau9p, and extension of Pau21/22p (Fig. 4a), might be due to sequencing errors. To test this, we amplified and sequenced *PAU7*, *PAU9* and *PAU21/22* from an S288C-derivative strain, BY4742, and a wine yeast strain of *S. cerevisiae*, S92 (Fig. 4b). In BY4742, in the sequence of *PAU7*, the nucleotide at position 166 is thymidine (T), confirming the reference sequence. However, *PAU7* in the wine yeast strain S92 has a cytidine (C) rather than a thymidine (T) at position 166, extending Pau7p in S92 to 122 aa, the same size as Pau7p in the relatively ancient species *S. paradoxus* and *S. mikatae*. The *PAU9* gene amplified from both BY4742 and S92 has two Ts starting at the −180 position rather than three Ts as reported in the reference sequence (Fig. 4b), suggesting a sequencing error in *PAU9* in the reference sequence. As a consequence, the corresponding Pau9p is corrected to 120 aa long, thus sharing an identical

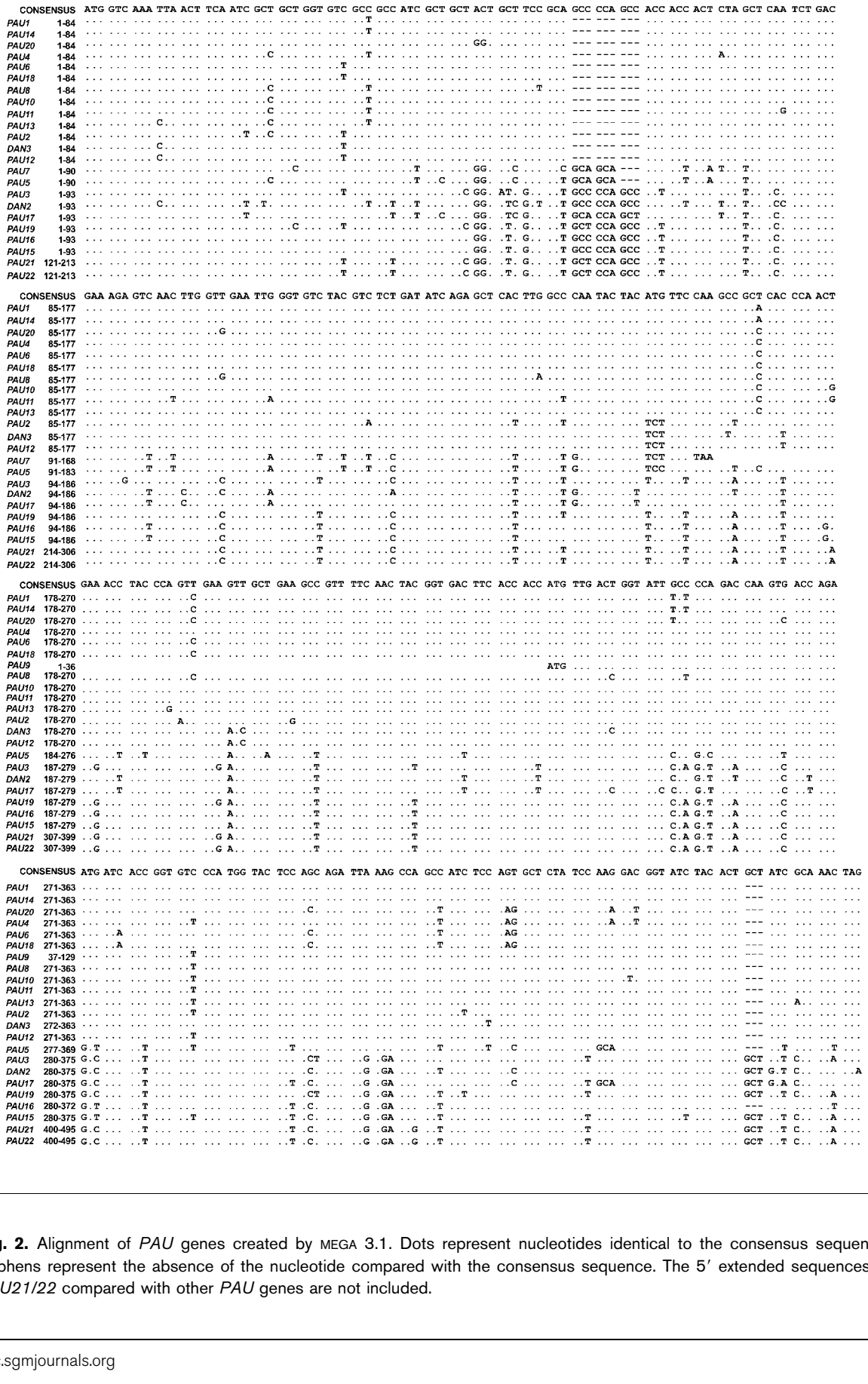


Fig. 2. Alignment of *PAU* genes created by MEGA 3.1. Dots represent nucleotides identical to the consensus sequence. Hyphens represent the absence of the nucleotide compared with the consensus sequence. The 5′ extended sequences of *PAU21/22* compared with other *PAU* genes are not included.

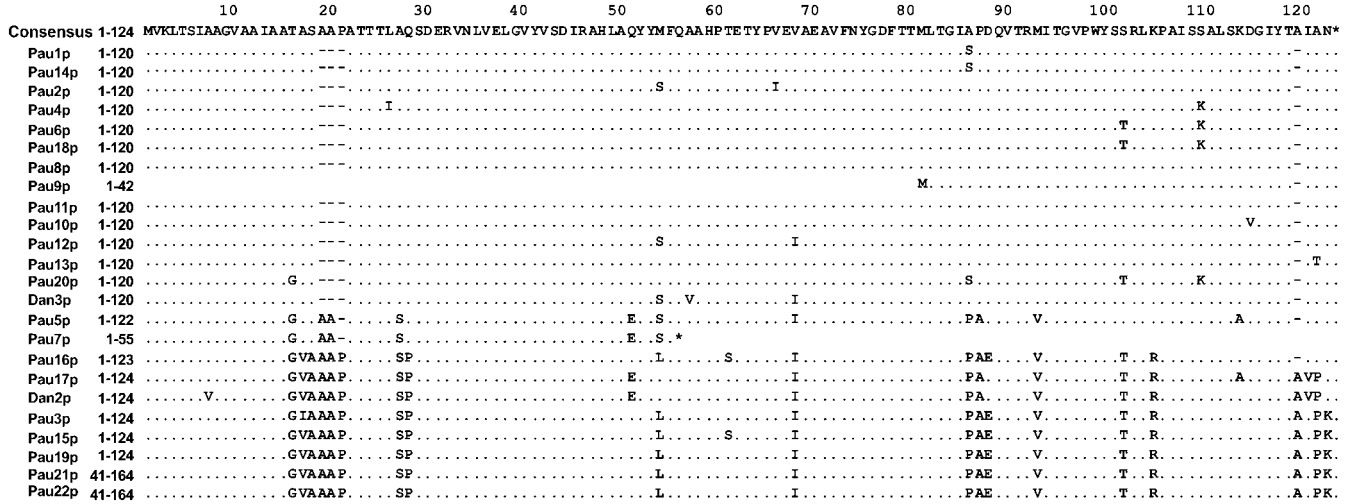


Fig. 3. Alignment of Pau proteins created by MEGA 3.1. Dots represent amino acids identical to the consensus sequence. Hyphens represent the absence of amino acids compared with the consensus sequence. The N-terminal extended sequences of Pau21/22p compared with other Pau proteins are not included.

sequence with Pau8p and Pau11p. These sequencing results were further confirmed by Western blot analyses showing that the molecular masses of GFP–Pau7p and GFP–Pau9 are almost the same as that of GFP–Pau5p (Fig. 4c). The sequences of *PAU21/22* in both BY4742 and S92 are the same as the reference sequences (Fig. 4b).

PAU21/22 encodes two peptides

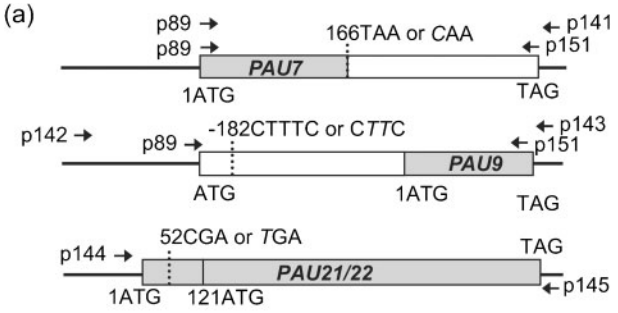
Pau21/22p contains a 40 aa N-terminal extension compared with other Pau proteins. The *PAU21/22* ORF contains three in-frame ATGs, with the second one corresponding to the translational start site of other *PAU* genes (Fig. 4a). We speculated that *PAU21/22* produced two peptides of different lengths by translating from different start codons. To test this possibility, we cloned *PAU21/22* with their promoter sequence, changed two potential start codons, Met1 and Met41, respectively, to isoleucine, and constructed C-terminal TAP tag-fused expression cassettes in a 2μ plasmid pRS326. As *PAU* genes have been reported to be anaerobic genes (Rachidi *et al.*, 2000), we tested the expression of Pau21/22p in anaerobically grown cells. Western blot analysis showed that under the anaerobic conditions, cells expressed wild-type Pau21/22p-TAP as two major proteins (36 and 40 kDa) (Fig. 4d, lane 1), mutation of the first ATG of *PAU21/22* resulted in the disappearance of the 40 kDa band (Fig. 4d, lane 2), and mutation of the second ATG at position 121 made the 36 kDa band disappear (Fig. 4d, lane 3), demonstrating that ATGs at both position 1 and position 121 in *PAU21/22* are functional and initiate the translation of peptides of 164 and 124 aa, respectively. From the relative amount of the two peptides, the small one of similar size to other Pau proteins is the major form induced by anaerobiosis. Whether the 40 aa extension of Pau21/22p at its N terminal is function-related remains unknown.

Relationship of *PAU* genes and their relatives

Pau proteins share homology with the N-terminal region of Tir and Dan proteins. To establish their genetic relationship, a phylogenetic tree was constructed using MEGA 3.1 for all *PAU*, *DAN* and *TIR* genes in *S. cerevisiae* (Fig. 5). Our data indicated that *PAU* genes are clustered into three groups that encode proteins of 120, 122 and 123/124 aa. The *PAU* gene family is more closely related to the *DAN* family than to the *TIR* family. *DAN2* was clustered in group III of the *PAU* family and *DAN3* in group I of the *PAU* family; they were obviously misnamed previously (Abramova *et al.*, 2001a). We therefore suggest that *DAN2* and *DAN3* be renamed as *PAU23* and *PAU24*, respectively.

Comparison of flanking sequences among *PAU* genes

To determine potential divergence in the flanking regions of the *PAU* genes, we retrieved and compared up to 500 bp upstream of the start codon and 500 bp downstream of the stop codon from all 24 *PAU* genes (Table 3). In general, the similarity varies greatly between the flanking sequences of the *PAU* genes: from no significant similarity to 100 % identity. Most of the subtelomeric *PAU* genes showed high homology in the flanking regions, while the internal *PAU* genes (*PAU2*, *PAU5*, *PAU7*, *PAU17* and *DAN2*) showed no significant similarity, except for *PAU5* and *PAU7*, which had 79 % identity in their 5′ 500 bp flanking sequence. Among the higher-homology flanking sequences in all *PAU* genes, the minimum identity was 66 %, lower than that of



(b)

ORFs		*PAU7*	*PAU9*	*PAU21/22*
Sequences from this study	BY4742	166TAA	-182CTTC	52CGA
	S92	166CAA	-182CTTC	52CGA
Reference sequences	S288C	166TAA	-182CTTTC	52CGA

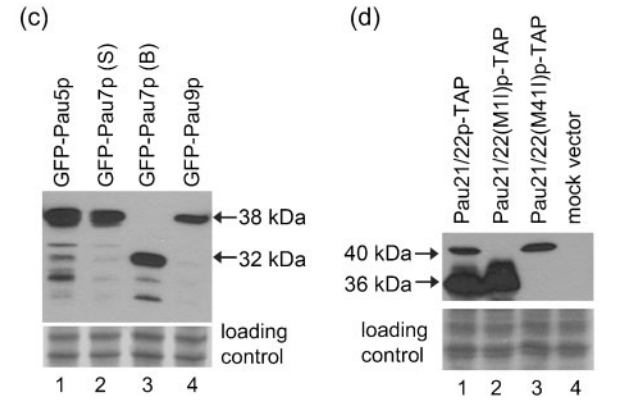


Fig. 4. Sequencing, cloning and expression of *PAU7*, *PAU9* and *PAU21/PAU22*. (a) Schematic representation of ORFs for *PAU7*, *PAU9* and *PAU21/PAU22*. The dotted line indicates the position at which the nucleotide change will lead to the extension of the *PAU7* ORF at the 3′ end, the *PAU9* ORF at the 5′ end, or truncation of the *PAU21/PAU22* ORF at the 5′ end to a size similar to that of other *PAU* ORFs. Arrows indicate the position and direction of oligonucleotides used for PCR amplification, sequencing and cloning. The number in front of the nucleotide indicates its position in the ORF. (b) Summary of the sequencing results. (c) Immunoblotting of cell lysates from cells expressing GFP–Pau fusions in a single-copy plasmid driven by the *MET25* promoter. 'S' indicates that the *PAU7* sequence was from strain S92; 'B' indicates that the *PAU7* sequence was from strain BY4742. (d) Immunoblotting of cell lysates from cells expressing wild-type and two mutant Pau21/22p–TAP fusion proteins driven by their native promoters on a multicopy plasmid. Cells were grown under anaerobic conditions.

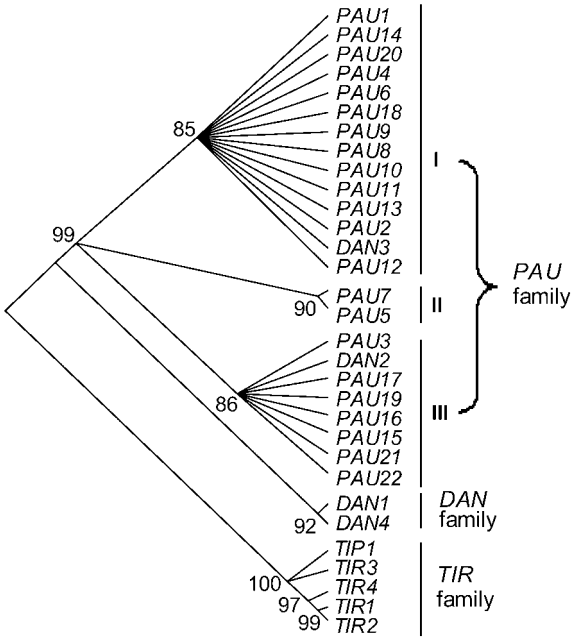


Fig. 5. Phylogenetic relationships of *PAU*, *DAN* and *TIR* genes. The phylogenetic tree (neighbour-joining method) was produced with MEGA 3.1. The bootstrap test of phylogenies was performed in 1000 replicates. The cut-off value was set at 85.

the coding sequence (85 %), suggesting that flanking sequences have diverged faster than the coding sequences.

Duplicated genes can result from large-scale duplication, e.g. chromosomal or segmental duplication. In this type of duplication, the duplicated genes show high similarity in both the coding regions and their flanking regions. The duplication of the most subtelomeric *PAU* genes might belong to this type. For example, *PAU21* and *PAU22* share 100 % identity in the coding regions and 99 % in the upstream regions (up to a few kilobases) and downstream regions (500 bp). The gene order in the *PAU21* and *PAU22* neighbourhood is *PAU21-ERR1-HSP33—FDH1* and *PAU22-ERR2-HSP32—FDH2*, respectively. It is obvious that the gene pair of *PAU21* and *PAU22* is located within the segment that was duplicated between the right end of chromosome XV and the left end of chromosome XVI. The length of this duplication spans at least 11 146 kb. The other type of duplication is called retroposition (Zhang, 2003). This type occurs when an mRNA is retrotranscribed to cDNA, then inserted into the genome. Thus the promoter and termination sequences of the resulting duplicates usually have a different origin. Some of the internal *PAU* genes or the other *PAU* genes whose flanking sequences lack similarity might have originated through retroposition.

Motif searching using the YEASTRACT database (Teixeira *et al.*, 2006) revealed some common features in the promoters of *PAU* genes. All but one contain at least one copy of the Upc2p binding site (absent in *PAU13*) and the

Table 3. Sequence similarity in 500 bp flanking regions between *PAU* genes

The non-shadowed numbers represent the percentage identity between 500 bp sequences upstream of the start codon of *PAU* genes; the shadowed numbers represent the percentage identity between 500 bp sequences downstream of the stop codon of *PAU* genes. –, No significant similarity.

Identity (%)	PAU1	PAU2	PAU3	PAU4	PAU5	PAU6	PAU7	PAU8	PAU9	PAU10	PAU11	PAU12	PAU13	PAU14	PAU15	PAU16	PAU17	PAU18	PAU19	PAU20	PAU21	PAU22	DAN2	DAN3
PAU1	100	–	–	92	–	93	–	94	91	94	–	–	–	99	86	95	–	93	74	95	74	74	–	–
PAU2	–	100	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
PAU3	92	–	100	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
PAU4	92	–	91	100	–	94	–	91	92	92	–	–	–	92	–	95	–	94	79	93	75	75	–	–
PAU5	–	–	–	–	100	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
PAU6	93	–	94	88	–	100	–	92	95	91	–	–	–	93	–	94	–	99	73	95	74	74	–	–
PAU7	–	–	–	–	79	–	100	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
PAU8	88	–	88	88	–	85	–	100	91	93	–	–	–	94	–	94	–	92	76	91	75	75	–	–
PAU9	90	–	95	90	–	92	–	85	100	94	–	–	–	91	–	92	–	94	77	94	73	73	–	–
PAU10	67	–	66	70	–	66	–	71	68	100	–	–	–	94	–	94	–	91	75	91	74	74	–	–
PAU11	92	–	90	91	–	90	–	88	89	68	100	–	–	–	–	–	–	–	–	–	–	–	–	–
PAU12	–	–	–	–	–	–	–	–	–	–	–	100	–	–	–	–	–	–	–	–	–	–	–	96
PAU13	87	–	85	86	–	84	–	84	86	67	85	–	100	–	–	–	–	–	–	–	–	–	–	–
PAU14	100	–	92	92	–	93	–	88	90	66	92	–	87	100	–	96	–	93	74	95	75	75	–	–
PAU15	93	–	90	92	–	89	–	90	88	67	96	–	86	93	100	–	–	–	–	–	–	–	–	–
PAU16	92	–	89	92	–	90	–	90	88	67	95	–	86	92	98	100	–	94	76	94	73	73	–	–
PAU17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	100	–	–	–	–	–	–	–
PAU18	91	–	96	90	–	93	–	87	94	67	92	–	85	91	90	90	–	100	73	95	74	74	–	–
PAU19	84	–	80	82	–	81	–	80	80	68	82	–	86	84	83	83	–	79	100	76	94	94	–	–
PAU20	80	–	77	80	–	78	–	76	77	66	80	–	78	80	82	82	–	76	78	100	76	76	–	–
PAU21	84	–	80	83	–	81	–	81	80	68	82	–	85	84	83	83	–	80	95	77	100	99	–	–
PAU22	84	–	80	83	–	81	–	81	80	68	82	–	85	84	83	83	–	80	95	77	100	100	–	–
DAN2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	100	–
DAN3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	100

Mot3p binding site (absent in *PAU9*) (Table 4). Upc2p induces transcription of sterol biosynthetic genes (Vik & Rine, 2001) and of *DAN/TIR* genes (Abramova *et al.*, 2001b) under hypoxic/anaerobic conditions. Its binding motif (TCGTTYAG) is also called an anaerobic response element (AR1) (Cohen *et al.*, 2001). Mot3p is a transcription factor involved in repression of a subset of hypoxic genes by Rox1p, repression of several *DAN/TIR* genes during aerobic growth, and repression of ergosterol biosynthetic genes (Hongay *et al.*, 2002; Sertil *et al.*, 2003). Probably both Upc2p and Mot3p are also involved in the regulation of *PAU* gene expression. The other common feature in the *PAU* promoters is they all have a TATA box in the region of −95 to −110 before the start codon. These findings are in agreement with a previous report that TATA-containing promoters are prevalent in subtelomeric regions, highly associated with response to stresses, and highly evolving (Basehoar *et al.*, 2004).

Differential expression of *PAU* genes

Duplicated genes may gain novel functions through regulatory evolution. We therefore investigated whether the expression profiles of *PAU* genes are different. Given the high sequence similarity amongst the *PAU* genes, which makes it impossible to detect signals from individual *PAU* mRNAs due to cross-hybridization, we applied Western blot analyses to detect Pau protein levels. Five *PAU* genes, *PAU3*, *PAU5*, *DAN2*, *PAU17* and *PAU20*, representing different subgroups and chromosomal localizations, were fused with a TAP tag just before the stop codon in their chromosomal loci. The expression of Pau5p–TAP, which has been previously reported (Luo & van Vuuren, 2008), was used as a control. Western blot analysis revealed that like Pau5p–TAP, Dan2p–TAP and Pau20p–TAP were highly induced by low temperature and anaerobic conditions. Pau3p–TAP and Pau17p–TAP were not detected under any of the conditions tested (Fig. 6a). Since the stability of Pau proteins might differ, thus affecting steady-state levels, we measured the promoter activity by using the TAP tag as a reporter, replacing the *PAU* ORF with a TAP tag in the genome of BY4742. Expression was detected by Western blotting. As shown in Fig. 6(b), the basal constitutive expression of *PAU5* and *DAN2* was detected in cells grown in YPD. Treatment of yeast cells for 2 h with nitrogen starvation, carbohydrate starvation or growth in a

Table 4. Common DNA motifs in the *PAU* gene promoters

Each number represents the nucleotide position of the motifs upstream of the start codon of each *PAU* gene; 'c' after the number indicates that the binding motif is on the complementary strand.

Gene name	Anaerobic induction/Upc2p binding sites		Aerobic repression/Mot3p binding sites			TATA box
	TCGTATA	TCGTTYAG	TMGGAA	WAGGKA	CAGGYA	TATAAATA
PAU1	−526c		−723, −817c			−100
PAU2	−215c, −254c	−375c	−34c	−714c, −538c		−95
PAU3	−529c	−234, −419c	−708			−100
PAU4	−522c	−233	−718, −812c	−343		−100
PAU5	−244c, −743c			−942c	−690	−108
PAU6	−530	−417c	−718, −813c			−101
PAU7	−980, −242c		−247c, −734c			−108
PAU8	−523c	−230	−916, −905	−901	−362	−100
PAU9	−776c	−469, −654c				−101
PAU10	−253c	−434	−38c, −390c	−861, −510	−364	−109
PAU11	−525c	−233	−927c	−921	−364	−100
PAU12	−262, −530c	−424c	−344	−290, −354c	−875	−101
PAU13			−823c	−50c, −270c		−101
PAU14	−526c		−723, −817c			−100
PAU15	−525c	−233	−928c	−922	−364	−100
PAU16	−525c	−233			−364	−100
PAU17	−964c	−211, −409c	−629c	−339	−946c	−107
PAU18	−543c	−234	−508			−109
PAU19	−535c	−465, −416c		−104, −259c		−100
PAU20	−526c	−463, −232	−721	−261c	−369	−100
PAU21	−418c	−347, −298c	−970	−141c, −401c	−245	−100
PAU22	−418c	−347, −298c	−970	−141c, −401c	−245	−100
DAN2	−263c	−447	−76c	−353, −296	−554	−111
DAN3	−262, −529c		−343	−353, −290	−874	−101

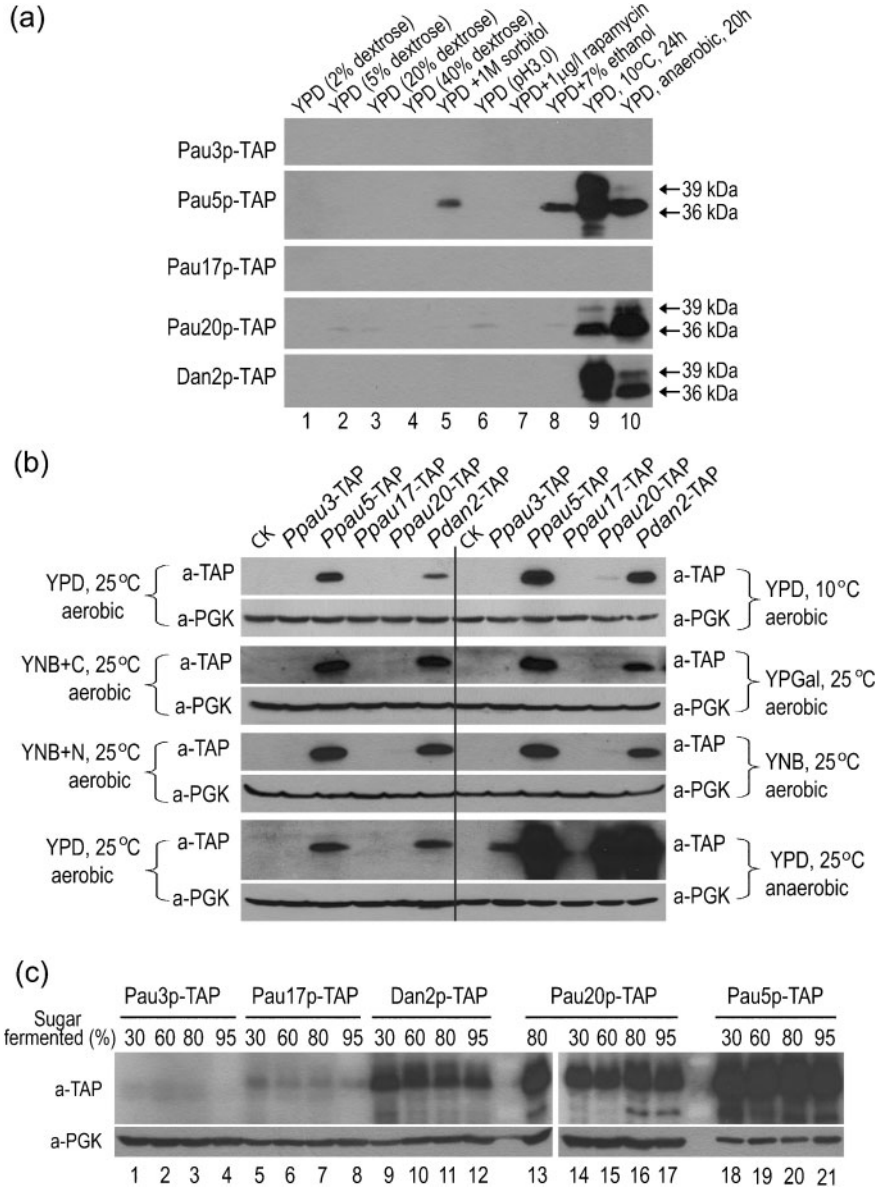


Fig. 6. Differential expression of *PAU* genes. (a) Immunoblotting of cell lysates from strains LY5, LY6, LY7, LY8 and LY9, which chromosomally encode Pau3p–TAP, Pau5p–TAP, Dan2p–TAP, Pau17p–TAP and Pau20p–TAP, respectively. Growth media and conditions were as indicated. (b) Immunoblotting of cell lysates from strains LY19, LY20, LY280, LY281 and LY282 (chromosomally encoding a TAP tag under the control of the native promoter of *PAU3*, *PAU5*, *DAN2*, *PAU17* and *PAU20*, respectively). Aerobically grown exponential phase cells were collected, washed and resuspended in the media and grown under the conditions as indicated. YNB, yeast nitrogen base without ammonium sulphate and amino acids. Loading controls were visualized by immunoblotting the blots with anti-PGK mAbs. (c) Immunoblotting of cell lysates from wine yeast strains LY12-7, LY15-8, LY21-7, LY22-8 and LY23-8 (chromosomally encoding Pau3p–TAP, Pau5p–TAP, Dan2p–TAP, Pau17p–TAP and Pau20p–TAP, respectively). Cells were collected at various stages during fermentation of Chardonnay grape must. Loading controls were visualized by immunoblotting with anti-PGK mAbs.

non-fermentative medium showed the least effect on the induction of the five *PAU* genes. Under low temperature, the induction of TAP was observed in *PAU5*, *PAU20* and *DAN2*, with much more induction of *PAU5* and *DAN2*. Under anaerobic conditions, activity of the *PAU3* promoter was weakly induced, while the activity of the *PAU5*, *PAU20* and *DAN2* promoters was dramatically increased. The induction of Pau–TAP fusion proteins in wine yeast strains was also investigated during alcoholic fermentation using Chardonnay grape must. The results (Fig. 6c) showed

that Pau5p–TAP, Pau20p–TAP and Dan2p–TAP were largely induced at an early stage of fermentation and remained almost constant throughout the entire fermentation process. Pau17p–TAP was lightly induced, while almost no Pau3–TAP was observed. Obviously, these five *PAU* genes responded differently to stress.

The repression of *PAU* genes under aerobic conditions has been suggested to be caused by the telomere position effect (TPE) (Ai *et al.*, 2002). To determine whether this is indeed the case, we studied the expression of the *URA3* gene by growing Pau–TAP-tagged yeast strains on 5-fluoro-orotic acid (5-FOA) plates. In Pau–TAP-tagged strains, a copy of the *URA3* gene with its native promoter is located in the region immediately downstream of the *PAU* genes. If *URA3* is expressed, it would make 5-FOA toxic and yeast cells would not grow (Boeke *et al.*, 1984). Our results showed that as control, the strain BY4742 (*ura3Δ0*) grew well on the 5-FOA plate, while strains LY5 (*PAU3-TAP-URA3)*, LY6 (*PAU5-TAP-URA3*), LY7 (*DAN2-TAP-URA3*), LY8 (*PAU17-TAP-URA3*) and LY9 (*PAU20-TAP-URA3*) all failed to grow on the 5-FOA plate (data not shown), indicating that the native *URA3* gene located in a similar position to the *PAU* genes was expressed in all the *PAU-TAP-URA3* strains. Therefore, repression of the subtelomeric *PAU* genes might not be due to TPE.

A hypothesis has been proposed elsewhere that subtelomeric *PAU* genes are more weakly inducible than the internal ones, based on the expression of *PAU4*, *PAU5* and *PAU6* on the plasmid-borne copies of *PAU–lacZ* fusions (Rachidi *et al.*, 2000). Our results (Fig. 6), however, indicate that *PAU20*, although located in the subtelomeric region, was highly induced by anaerobic conditions and weakly induced by low temperature. Therefore, gene position does not appear to correlate to expression under different conditions. Instead, we believe that the variation in promoter sequences, the number of copies and location of the stress-inducible elements and the repressor binding motifs (Table 4), as well as the 5′ and 3′ untranslated regions (UTRs), might be responsible for the differential expression of *PAU* genes in response to stresses.

Divergence in stability of Pau proteins

To test whether the sequence divergence of Pau proteins affects their stability, we cloned all *PAU* ORFs into a *CEN*-based expression plasmid driven by the *MET25* promoter. Pau–TAP fusion proteins were induced by growing cells in YNB medium without methionine. Western blot analysis showed that after 12 h induction, levels of individual Pau proteins varied, with the average amount of 120 and 122 aa Pau proteins being about eightfold more than that of 124 aa Pau proteins (Fig. 7a), suggesting that the latter are less stable. To confirm this, cycloheximide chase assays were performed for Pau1p (120 aa) and Pau3p (124 aa). After protein synthesis was stopped by cycloheximide, Pau3–TAP degraded rapidly in YPD, with a half-life of less than 15 min, while the half-life of Pau1–TAP was ~30 min (Fig. 7b). Anaerobic growth conditions slowed the degradation process of both proteins. In addition, the molecular size of Pau1–TAP and Pau3–TAP gradually increased during their degradation process, in agreement with the previous observations for Pau5p–TAP (Luo & van Vuuren, 2008), which might be due to hyperphosphorylation. These results suggest that in addition to transcriptional regulation, yeast cells have evolved mechanisms to independently regulate the protein levels of the multiple *PAU* genes.

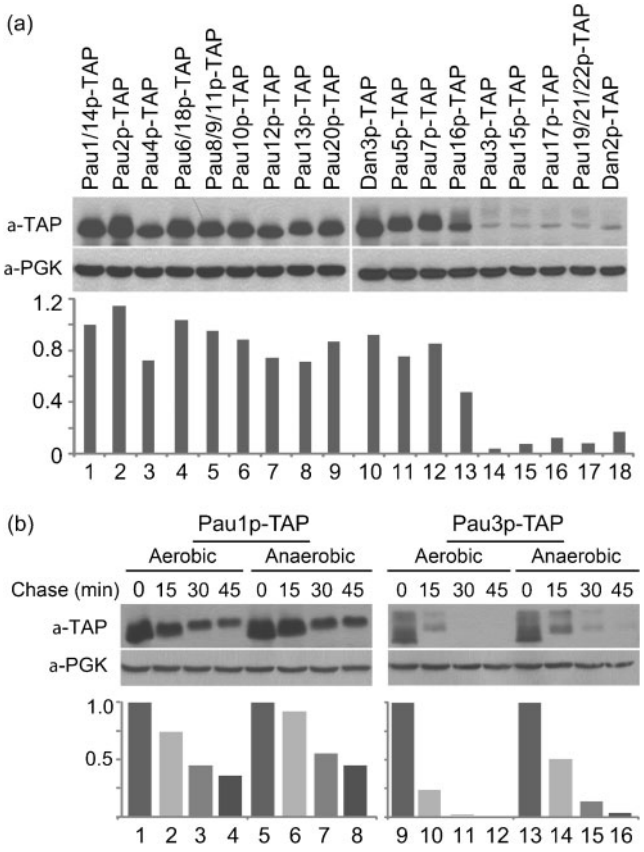


Fig. 7. Pau protein levels and their degradation. (a) Immunoblotting of cell extracts from exponentially growing cells that contained *CEN*-based plasmids encoding different Pau–TAP fusion proteins driven by the *MET25* promoter. Cells were grown in YNB plus 2 % glucose, ammonium sulphate and essential amino acids. (b) Immunoblotting of cell lysates from cycloheximide-treated cells. Cells expressing Pau1p–TAP and Pau3p–TAP under the control of the *MET25* promoter in a *CEN* plasmid were aerobically or anaerobically grown to mid-exponential phase in YNB plus 2 % glucose, ammonium sulphate and essential amino acids, exposed to cycloheximide, and chased under aerobic or anaerobic conditions for the times indicated. The blots were probed with anti-PGK antibodies as a reference. Charts at the bottom of (a) and (b) show the relative level of Pau5p–TAP normalized to the respective Pgk1p band signal. Signal quantification was performed with Image J (http://rsb.info.nih.gov/ij/).

In conclusion, this study provides evidence that after duplication, paralogues of the multiple *PAU* gene family have diverged at the levels of sequence, expression and protein stability. Although the function of Pau proteins has yet to be defined, the amplification and persistence of so many diversified *PAU* genes might indicate that they have specific roles in the adaptation of yeasts to a variety of environmental stresses.

ACKNOWLEDGEMENTS

This work was supported by an Natural Sciences and Engineering Research Council (NSERC) grant (217271-04) to H. J. J. v. V. We thank Christopher Walkey for his critical reading of this manuscript.

REFERENCES

Abramova, N. E., Sertil, O., Mehta, S. & Lowry, C. V. (2001a). Reciprocal regulation of anaerobic and aerobic cell wall mannoprotein gene expression in *Saccharomyces cerevisiae*. *J Bacteriol* **183**, 2881–2887.

Abramova, N. E., Cohen, B. D., Sertil, O., Kapoor, R., Davies, K. J. & Lowry, C. V. (2001b). Regulatory mechanisms controlling expression of the *DAN/TIR* mannoprotein genes during anaerobic remodelling of the cell wall in *Saccharomyces cerevisiae*. *Genetics* **157**, 1169–1177.

Ai, W., Bertram, P. G., Tsang, C. K., Chan, T. F. & Zheng, X. F. (2002). Regulation of subtelomeric silencing during stress response. *Mol Cell* **10**, 1295–1305.

Alimardani, P., Regnacq, M., Moreau-Vauzelle, C., Ferreira, T., Rossignol, T., Blondin, B. & Berges, T. (2004). *SUT1*-promoted sterol uptake involves the ABC transporter Aus1 and the mannoprotein Dan1 whose synergistic action is sufficient for this process. *Biochem J* **381**, 195–202.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.

Ausubel, F. M., Brent, R., Kinston, R. E., Moore, D. D., Scidman, J. D., Smith, J. A. & Struhl, K. (1999). *Short Protocols in Molecular Biology*, 4th edn. New York: Wiley.

Basehoar, A. D., Zanton, S. J. & Pugh, B. F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**, 699–709.

Blandin, G., Durrens, P., Tekaia, F., Aigle, M., Bolotin-Fukuhara, M., Bon, E., Casaregola, S., de Montigny, J., Gaillardin, C. & other authors (2000). Genomic exploration of the hemiascomycetous yeasts: 4. The genome of *Saccharomyces cerevisiae* revisited. *FEBS Lett* **487**, 31–36.

Boeke, J. D., LaCroute, F. & Fink, G. R. (1984). A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet* **197**, 345–346.

Byrne, K. P. & Wolfe, K. H. (2006). Visualizing syntenic relationships among the hemiascomycetes with the yeast gene order browser. *Nucleic Acids Res* **34**, D452–D455.

Cohen, B. D., Sertil, O., Abramova, N. E., Davies, K. J. & Lowry, C. V. (2001). Induction and repression of *DAN1* and the family of anaerobic mannoprotein genes in *Saccharomyces cerevisiae* occurs through a complex array of regulatory sites. *Nucleic Acids Res* **29**, 799–808.

Davies, B. S. & Rine, J. (2006). A role for sterol levels in oxygen sensing in *Saccharomyces cerevisiae*. *Genetics* **174**, 191–201.

Despons, L., Wirth, B., Louis, V. L., Potier, S. & Souciet, J. L. (2006). An evolutionary scenario for one of the largest yeast gene families. *Trends Genet* **22**, 10–15.

Fabre, E., Muller, H., Therizols, P., Lafontaine, I., Dujon, B. & Fairhead, C. (2005). Comparative genomics in hemiascomycete yeasts: evolution of sex, silencing, and subtelomeres. *Mol Biol Evol* **22**, 856–873.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C. & other authors (1996). Life with 6000 genes. *Science* **274**, 546–567.

Gu, Z., Nicolae, D., Lu, H. H. & Li, W. H. (2002). Rapid divergence in expression between duplicate genes inferred from microarray data. *Trends Genet* **18**, 609–613.

Homma, T., Iwahashi, H. & Komatsu, Y. (2003). Yeast gene expression during growth at low temperature. *Cryobiology* **46**, 230–237.

Hongay, C., Jia, N., Bard, M. & Winston, F. (2002). Mot3 is a transcriptional repressor of ergosterol biosynthetic genes and is required for normal vacuolar function in *Saccharomyces cerevisiae*. *EMBO J* **21**, 4114–4124.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254.

Kong, H., Landherr, L. L., Frohlich, M. W., Leebens-Mack, J., Ma, H. & dePamphilis, C. W. (2007). Patterns of gene duplication in the plant *SKP1* gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. *Plant J* **50**, 873–885.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

Luo, Z. & van Vuuren, H. J. J. (2008). Stress-induced production, processing and stability of a seripauperin protein, Pau5p, in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **8**, 374–385.

Lynch, M. & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155.

Marks, V. D., Ho Sui, S. J., Erasmus, D., van der Merwe, G. K., Brumm, J., Wasserman, W. W., Bryan, J. & van Vuuren, H. J. J. (2008). Dynamics of the yeast transcriptome during wine fermentation reveals a novel fermentation stress response. *FEMS Yeast Res* **8**, 35–52.

Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.

Ohno, S. (1970). *Evolution by Gene Duplication*. New York: Springer.

Pan, D. & Zhang, L. (2008). Tandemly arrayed genes in vertebrate genomes. *Comp Funct Genomics* 545269.

Pryde, F. E., Gorham, H. C. & Louis, E. J. (1997). Chromosome ends: all the same under their caps. *Curr Opin Genet Dev* **7**, 822–828.

Rachidi, N., Martinez, M. J., Barre, P. & Blondin, B. (2000). *Saccharomyces cerevisiae PAU* genes are induced by anaerobiosis. *Mol Microbiol* **35**, 1421–1430.

Rossignol, T., Dulau, L., Julien, A. & Blondin, B. (2003). Genome-wide monitoring of wine yeast gene expression during alcoholic fermentation. *Yeast* **20**, 1369–1385.

Sahara, T., Goda, T. & Ohgiya, S. (2002). Comprehensive expression analysis of time-dependent genetic responses in yeast cells to low temperature. *J Biol Chem* **277**, 50015–50021.

Sertil, O., Kapoor, R., Cohen, B. D., Abramova, N. & Lowry, C. V. (2003). Synergistic repression of anaerobic genes by Mot3 and Rox1 in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **31**, 5831–5837.

Taylor, J. S. & Raes, J. (2004). Duplication and divergence: the evolution of new genes and old ideas. *Annu Rev Genet* **38**, 615–643.

Teixeira, M. C., Monteiro, P., Jain, P., Tenreiro, S. R., Fernandes, A. R., Mira, N. P., Alenquer, M., Freitas, A. T., Oliveira, A. L. & Sá-Correia, I. (2006). The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **34**, D446–D451.

Vik, A. & Rine, J. (2001). Upc2p and Ecm22p, dual regulators of sterol biosynthesis in *Saccharomyces cerevisiae*. *Mol Cell Biol* **21**, 6395–6405.

Viswanathan, M., Muthukumar, G., Cong, Y. S. & Lenard, J. (1994). Seripauperins of *Saccharomyces cerevisiae*: a new multigene family encoding serine-poor relatives of serine-rich proteins. *Gene* **148**, 149–153.

Wirth, B., Louis, V. L., Potier, S., Souciet, J. L. & Despons, L. (2005). Paleogenomics or the search for remnant duplicated copies of the yeast *DUP240* gene family in intergenic areas. *Mol Biol Evol* **22**, 1764–1771.

Wyrick, J. J., Holstege, F. C., Jennings, E. G., Causton, H. C., Shore, D., Grunstein, M., Lander, E. S. & Young, R. A. (1999). Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402**, 418–421.

Zhang, J. (2003). Evolution by gene duplication: an update. *Trends Ecol Evol* **18**, 292–298.

Edited by: K. Kuchler