INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Functional overlap in gene duplications in the yeast *Saccharomyces cerevisiae*

Jennifer G. Reich

A Thesis

In

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Science at
Concordia University
Montréal, Québec, Canada

November 2000

Abstract

Functional overlap in gene duplications in the yeast *"Saccharomyces cerevisiae"*

Jennifer G. Reich

The *Saccharomyces cerevisiae* genome contains many duplicated regions of DNA. The presence of extra copies of some genes can result in certain genes performing similar or identical functions as their duplicated counterparts. In such cases, this "functional overlap" is an obstacle in determining gene function by traditional single-gene inactivation studies, as one inactivated copy of a gene may result in no visible effects ("phenotypes") to the organism in the presence of a functional copy. This may contribute to the approximate half of yeast genes whose protein functions remain unknown. This study reveals phenotypes for functionally overlapping genes and classifies selected duplicated genes as functionally overlapping or divergent.

A region of DNA on chromosomes I and XV containing 12 duplicated genes was examined. Mutants were created containing deletions of one or both members of every gene pair and subjected to several functional tests. Gene pairs were classified as functionally overlapping when both members had to be deleted

to observe a "synthetic" phenotype. Gene pairs were considered to be functionally divergent when the removal of one member of a gene pair revealed a mutant phenotype, as the remaining member was unable to "rescue" the phenotype.

Out of the 12 gene pairs studied, three show functional divergence: *OAF1/PIP2, PYK1/PYK2,* and *MYO4/MYO2*. Four gene pairs show functional overlap from examining double deletions: *YAL056w/YOR371c*, *YAL031c/YOR334w(MRS2)*, *SNC1/SNC2*, and *PMT2/PMT3*. Two more gene pairs, *YAL053w/YOR365c* and *GDH1/GDH3* show functional overlap between one or both paralogs and additional homologs outside this duplicated region. Supporting this, two mutants deleted for three homologs, *Δgdh1 Δgdh3 Δglt1* (Avendano et al. 1997) and *Δyal053w Δygl139w Δypl221w*, show synthetic phenotypes. For the remaining three gene pairs, *YAL037w/YOR342c*, *YAL034c/YOR338w*, and *YAL028w/YOR324c*, no mutant phenotypes were detected.

Acknowledgements

I offer my sincere gratitude to Dr. Storms for giving me the opportunity to study in his lab, and for his guidance. I am indebted to Dr. Peter Ulycznyj for being so helpful throughout my project. My colleagues Edith Munro, Yun Zheng and Annie Boutet have also contributed to this project, and have made this lab a very pleasant place to work in. I thank Dr. Dayanandan for generously offering his input in the phylogeny section. I also thank Dr. William Zerges and Dr. Adrian Tsang for their participation on my committee.

There are no words that can describe how fortunate I feel to have parents who have always believed in me. Any successes I achieve are in a large part theirs, too.

Dedication

To my parents, Julie and Lionel

Table of Contents

Table of Contents (ctd.)

List of Figures

List of Figures (ctd.)

List of Tables

1. INTRODUCTION

1.1 The *S. cerevisiae* genome

The 6,000 genes that comprise the yeast genome are of interest to researchers because of the parallels that can be drawn between gene function in this basic eukaryotic cell, and other more complex cells. For example, it is estimated that 20% of human proteins have orthologs (genes in different species related by vertical descent from a common ancestor) in yeast (Mushegian et al. 1998). The simplicity of the life cycle of the yeast cell and its ability to be quickly and easily manipulated in the laboratory has resulted in its widespread use as a model organism for research aimed at determining eukaryotic cell function.

Currently, the functions for slightly over half of all yeast proteins are known (www.mips.biochem.mpg.de/proj/yeast; Mewes et al. 1999); (Fig. 1). Of the remaining half, 3.7% of sequences show strong similarity to known proteins. Proteins with some similarity to known proteins, unknown proteins, or no similarity to any known proteins each comprise 12-13% of the putative proteins encoded by the remaining open reading frames (ORFs).

It is believed that a large proportion of the genes in *S. cerevisiae* remain uncharacterized because they "escape" functional identification by current research methods. For example, Goebl and Petes (1986) determined that 70% of gene disruptions produce no obvious growth phenotypes. Several reasons have been proposed to account for the low proportion of genes that display a mutant phenotype when inactivated. First, laboratory tests probably do not replicate some

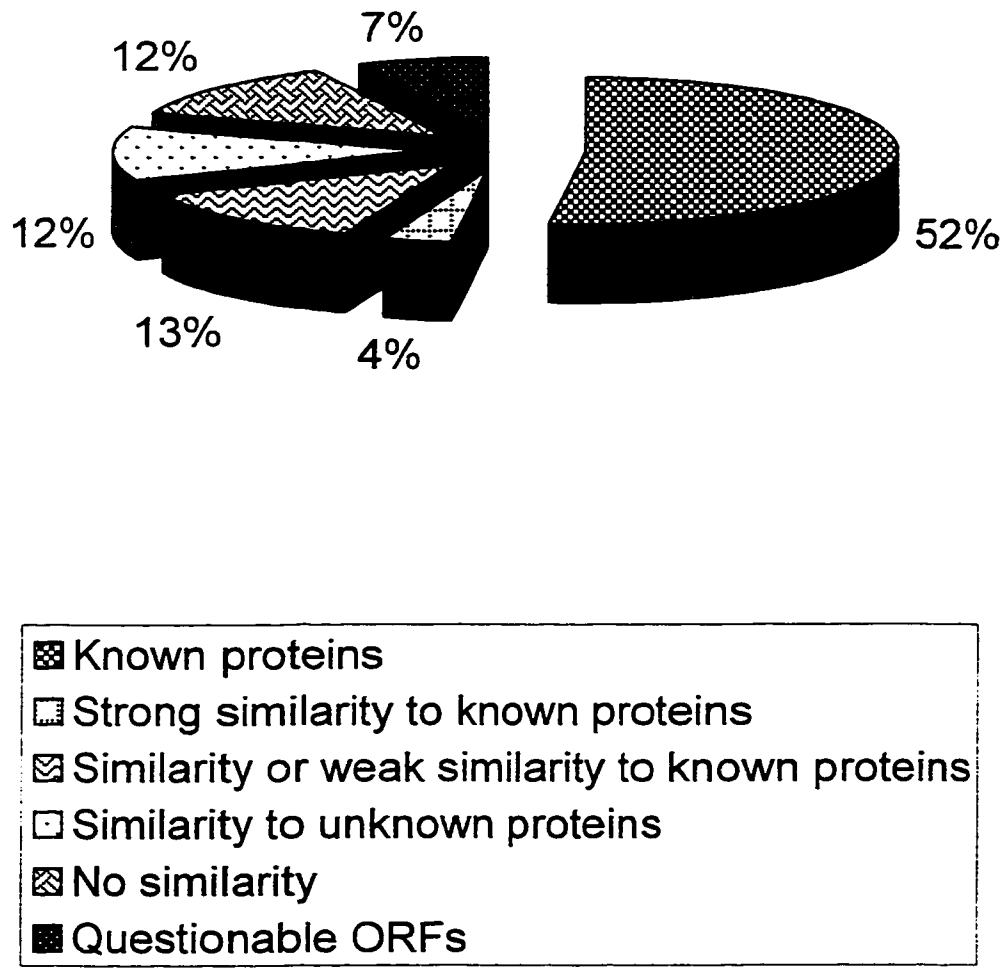


Fig. 1. Classification of yeast ORFs. MIPS (Mewes et al. 1999) analyzed 6362 putative open reading frames and categorized their encoded proteins as shown.

of the environmental conditions under which some genes are functional (Oliver, 1996). Second, the number of coding ORFs may be significantly lower than previously estimated (4700-4800 instead of 5800-6000). Mackiewicz et al. (1999) suggest that the "orphan paradox" (the higher than expected proportion of genes with unknown functions and without homologs) occurs when non-coding ORFs inside coding sequences in non-coding frames are transferred into other regions by duplication mechanisms. Third, some genes may make small but significant contributions to fitness that are not large enough to be detected by conventional methods (Thatcher et al. 1998). Finally, functionally redundant pathways and structures account for a proportion of the genes that lack an identifiable phenotype.

In cases of functional redundancy, analogous genes (unrelated genes performing the same function) and paralogous genes (genes in the same species related by duplication) functionally replace one another. Consequently, strains bearing single gene deletions may lack a mutant phenotype since their functions are duplicated elsewhere in the genome. For example, in a duplicated chromosomal region between nucleotides 29 000 and 106 000 on chromosome I and nucleotides 916 000 to 1 043 000 on chromosome XV, a significantly lower proportion of the chromosome I genes that had paralogs on chromosome XV were found to be functionally important as compared to the genes in this region which did not have homologous counterparts (Storms et al. 1997). That is, cell function was more often compromised when a gene that lacked a homologous partner was inactivated than when the mutated gene had a homologous counterpart. It was

postulated that functional overlap between the homologous gene pairs could be responsible for the lower incidence of phenotypes among chromosome I deletion mutants with homologous partners than among those without.

An overview of the sequenced yeast genome (Goffeau 1996) revealed that it contains many large regions of duplicated genes. The duplicated genes in each "block" are arranged in the same order on different chromosomes (Lalo et al. 1993; Melnick and Sherman, 1993; Goffeau et al. 1996, Coissac et al., 1997; Mewes et al. 1997; Philippsen et al; 1997; Wolfe and Shields 1997) and are interspersed with single copy genes. Smith (1987) originally suggested a polyploid origin for the 16 chromosome *S. cerevisiae* genome. Wolfe and Shields (1997) suggested that the *S. cerevisiae* diploid genome is a degenerate tetraploid resulting from a 10^8 year-old whole-genome duplication that occured after its divergence from *Kluyveromyces* 1.5 x 10^8 years ago. They believe that most duplicates were deleted and many reciprocal translocations occurred subsequent to the duplication event. Currently a mere 16% of the proteome remains as duplicates, many of which are located within the 53 such blocks (Seoighe and Wolfe 1999).

In contrast, *Kluyveromyces lactis* has only 6 chromosomes. Single copies of genes duplicated in *S. cerevisiae* are the norm (Wolfe and Shields, 1997) and gene order is the same as expected for a species that diverged from ancestors of *Saccharomyces* before the genome duplication (Seoighe and Wolfe 1999). Since *S. cerevisiae* has more than double the number of chromosomes found in *K. lactis,*

centromeric duplications resulting in new chromosomes are also likely to have occurred in this case (Wolfe & Lohan 1994).

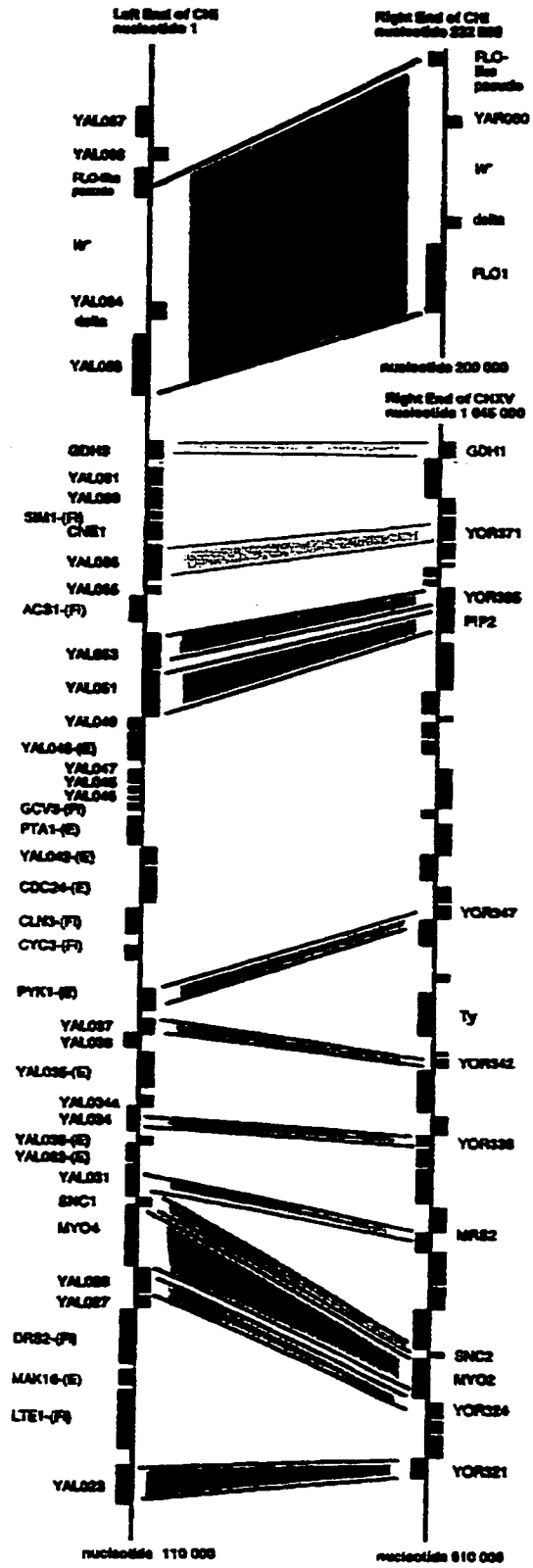
Gene duplication is a common and important evolutionary process in organisms (Ohno 1970), as it provides an opportunity for the acquisition of new gene functions and the diversification of physiological pathways. Whole-genome duplications result in having all regulatory as well as structural and enzymatic genes duplicated, thus maintaining gene-dosage relationships. As a result, all genes can evolve, including duplicated regulators which can co-evolve to differentially control their target genes.

1.2 Purpose of this study

Due to the presence of paralogous gene sets in *S. cerevisiae*, the issue of gene redundancy should be addressed in order to assign functional roles to the individual genes. Although single gene studies often take functional overlap into account when addressing gene function, to date a systematic analysis of duplicated genes has not been performed to address the issue of functional overlap.

The current study used single and double deletion mutants of the duplicated gene pairs associated with "block 2" (Fig. 2) on chromosomes I and XV (Wolfe and Shields, 1997) to assess whether some members of gene pairs exhibited functional overlap or were divergent. Each duplicated gene in "block 2" was singly replaced with a kanamycin resistance cassette, to create 24 deletion mutants. These

Fig. 2. Schematic diagram (Storms et al. 1997) of "block 2", a homologous region of DNA on the left end of chromosome I and the right end of chromosome XV where 12 gene duplications are found interspersed with single copy genes. Bars connect duplicate genes, and left/right orientation of black boxes represents the direction of gene transcription (towards/away from their telomeres, respectively).


Left End of ChII
nucleotide 1
Right End of ChII
YAL067
YAL066
FLO-like pseudo
W'
YAL064
delta
YAL060
FLO1
nucleotide 200 000
Right End of ChXV
nucleotide 1 045 000
GDH3
YAL061
YAL059
SIM1-(FI)
CNE1
YAL056
YAL055
ACS1-(FI)
YAL053
YAL051
YAL049
YAL048-(E)
YAL047
YAL045
YAL046
GCV3-(FI)
PTA1-(E)
YAL043-(E)
CDC24-(E)
CLN3-(FI)
CYC3-(FI)
PYK1-(E)
YAL037
YAL036
YAL035-(E)
YAL034
YAL034
YAL033-(E)
YAL032-(E)
YAL031
SNC1
MYO4
YAL028
YAL027
DRS2-(FI)
MAK16-(E)
LTE1-(FI)
YAL023
nucleotide 110 000
GDH1
YOR371
YOR365
PIP2
YOR347
Ty
YOR342
YOR338
MRS2
SNC2
MYO2
YOR324
YOR321
nucleotide 810 000

mutants were also used to construct strains deleted for both members of each syntenic gene pair. The resulting single and double deletion mutants were subjected to a series of functional tests. In one case where a strain with a doubly deleted "block 2" gene pair showed no mutant phenotype and had homologs on other chromosomes, mutants containing deletions of all combinations of homologs were created.

Cladistic analysis was used to determine "gene family" relationships for duplicates in "block 2" which had homologs on other chromosomes.

2. MATERIALS AND METHODS

2.1.1 Yeast strains and strain construction

The yeast strains used are listed in Table 1 and oligonucleotides used are listed in Table 2. Gene replacement was performed as described in Winzeler et al. (1999) using a modification of the PCR-based gene replacement method (Baudin et al. 1993). The ORF of each target gene was replaced by a KanMX4 cassette, which confers G418 resistance (Wach 1996). For each gene replacement, the KanMX4 cassette was used to replace the complete ORF of the gene of interest (Fig. 3). All strains containing deletions of chromosome I genes (except the *PMT2* deletion mutant) were created by first amplifying the cassette with 5' "uptag" and 3' "downstream" primers and then in a second round of PCR, by priming the amplified product with 5' "upstream" and 3' "downstream" primers. The first

Table 1. *S. cerevisiae* strains and plasmids used.

Strain No.	Genotype
5123	[1]*MAT*a *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5173-74	[1]*MAT*a *Δgdh1::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5175-76	[1]*MAT*a *Δyor371c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5177-78	[1]*MAT*a *Δyor365c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5179-80	[1]*MAT*a *Δ pip2c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5181-82	[1]*MAT*a *Δpyk2c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5183-84	[1]*MAT*a *Δyor342c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5185-86	[1]*MAT*a *Δyor338w::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5187-88	[1]*MAT*a *Δmrs2::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5189-90	[1]*MAT*a *Δsnc2::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5193-94	[1]*MAT*a *Δyor324c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5195-96	[1]*MAT*a *Δpmt3::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5307	[2]*MAT*∝ *Δypl221w::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5125	[3]*MAT*∝/*MAT*a *leu2Δ0/ leu2Δ0 MET15/met15Δ0 ura3Δ0/ ura3Δ0 his3Δ1/ his3Δ1 lys2Δ0/LYS2*
5171-72	[2]*MAT*∝ *Δpmt2:: kan*^{R} *leu2Δ0 ura3Δ0 his3Δ lys2Δ0*
5153-54	[2]*MAT*∝ *Δgdh3:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5155-56	[2]*MAT*∝ *Δyal056w:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5157-58	[2]*MAT*∝ *Δyal053w:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5159	[2]*MAT*∝ *Δoaf1:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5161-62	[2]*MAT*∝ *Δyal037w:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5163-64	[2]*MAT*∝ *Δyal034c:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5165-66	[2]*MAT*∝ *Δyal031c:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5167-68	[2]*MAT*∝ *Δsnc1:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5295	[2]*MAT*∝ *Δyal029:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5169-70	[2]*MAT*∝ *Δyal028:: kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5197	[1]*MAT*a *Δgdh3:: kan*^{R} *Δgdh1::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1 lys2Δ0*
5198	[1]*MAT*∝ *Δgdh3:: kan*^{R} *Δgdh1::kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5199	[1]*MAT*∝ *Δyor371c:: kan*^{R} *Δyal056w::kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*
5200	[1]*MAT*∝ *Δyor371c:: kan*^{R} *Δyal056w::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5201	[1]*MAT*a *Δyor365:: kan*^{R} *Δyal053::kan*^{R} *leu2Δ0 ura3Δ0 his3Δ1*
5202	[1]*MAT*a *Δ yor365:: kan*^{R} *Δyal053::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5203	[1]*MAT*a *Δpip2:: kan*^{R} *Δoaf1::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1 lys2Δ0*
5304	[4]*MAT*a *Δpyk2:: kan*^{R} *Δpyk1::kan*^{R} *leu2Δ0 ura3Δ0 his3Δ1*
5305	[4]*MAT*a *Δpyk2:: kan*^{R} *Δpyk1::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5205	[1]*MAT*a *Δyor342c:: kan*^{R} *Δyal037w::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*
5206	[1]*MAT*a *Δyor342c:: kan*^{R} *Δyal037w::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0*
5207	[1]*MAT*∝ *Δyor338w:: kan*^{R} *Δyal034c::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0 his3Δ1 lys2Δ0*
5208	[1]*MAT*∝ *Δyor338w:: kan*^{R} *Δyal034c::kan*^{R} *leu2Δ0 ura3Δ0*
5353	[4]*MAT*∝ *Δyor334:: kan*^{R} *Δyal031::kan*^{R}
5211	[1]*MAT*a *Δyor324c:: kan*^{R} *Δyal028w::kan*^{R} *leu2Δ0 met15Δ0 ura3Δ0*
5212	[1]*MAT*∝ *Δyor324c:: kan*^{R} *Δyal028w::kan*^{R} *leu2Δ0 ura3Δ0 lys2Δ0*

Table 1 (ctd.). *S. cerevisiae* strains and plasmids used.

Strain No.	Genotype
5313	[4]*MAT ∝ Δyal053w:: kan^R Δyor365:: kan^R Δypl221w:: kan^R leu2Δ0 met15Δ0 ura3Δ0 his3Δ1 LYS2*
5322	[5]*MATa Δygl139w::kan^R ura3-52 HIS3 LEU2 LYS2 trp1Δ63*
5328	[4]*MAT ∝ Δygl139w::kan^R Δyor365c:: kan^R Δypl221w:: kan^R*
5329	[4]*MAT ∝ Δygl139w::kan^R Δyor365c:: kan^R Δypl221w:: kan^R (His$^+$ Leu$^+$ Met$^+$ Lys$^+$ Trp$^-$ Ura$^-$)*
5330	[4]*MAT a Δygl139w::kan^R Δyor365c:: kan^R Δypl221w:: kan^R (His$^+$ Leu$^+$ Met$^-$ Lys$^+$ Trp$^+$ Ura$^-$)*
5333-35	[4]*MAT ∝ Δyal053w:: kan^R Δygl139w::kan^R Δyor365:: kan^R*
5336	[4]*MAT ∝ Δygl139w::kan^R Δypl221w:: kan^R*
5344-46	[4]*MAT a/∝ Δygl139w::kan^R Δyor365c::kan^R Δypl221w:: kan^R (His$^+$ Leu$^+$ Met$^-$ Lys$^+$ Trp$^-$ Ura$^-$)*
5347-48	[4]*MAT a/∝ Δygl139w::kan^R Δypl221w:: kan^R (His$^+$ Leu$^+$ Met$^+$ Lys$^+$ Trp$^+$ Ura$^-$)*
5349-51	[4]*MAT a/∝ Δyal053w::URA3/YAL053w Δygl139w::kan^R Δypl221w:: kan^R (His$^+$ Leu$^+$ Met$^+$ Lys$^+$ Trp$^-$ Ura$^+$)*
5352	[4]*MAT a/∝ Δyal053w::URA3/YAL053w Δygl139w::kan^R Δyor365c:: kan^R Δypl221w::kan^R (His$^+$ Leu$^+$ Met$^+$ Lys$^+$ Trp$^+$ Ura$^+$)*

Plasmids	Description	Reference
[3]KanMX4	contains Kanamycin cassette	Wach 1994.
[3]Yep352	contains URA3 gene	Hill et al. 1993.

Strains 5173-5196 and 5307 are derived from parent strain 5123 (BY4741). 5171 and 5172 are derived from 5125 (BY4743). Strains 5153-59, 5161-70 and 5295 are derived from [2]BY4739. Strains 5197-5203, 5205-5208, 5211-5212, 5304-5 and 5353 were derived by crossing BY4741 mutants to BY4739 mutants and sporulating the diploids. 5322 is derived from parent strain FY. Parent strains for 5313, 5328-30, 5333-34, 5344-52 are crosses between mutants created for this study (see text for details). Auxotrophic markers are indicated in parentheses where data is available.

Strains provided/created by: [1]Dr. Sophia Ushinsky (Concordia); [2]Dr. Howard Bussey and Steeve Veronneau (McGill); [3]Dr. R.K. Storms (Concordia); [4]This study; [5]Euroscarf.

Table 2. Oligonucleotides used for ORF replacement.

YAL062W:	[a]US:	AAGTAGCAACAGTCACCGAAAAGAAAAGGTAAAAAGTAAAAAATGGATGTCCACGAGCTCTCT
	[b]DS:	ATACACAGATAGTTACGAACAAAAAGAAAATAGCGCTTACGGCTAATCGATGAATTCGAGCTCG
	[c]UT:	GATGTCCACGAGGTCTCTCAGTATTGGGAATCGCTATCCGTACGCTGCAGGTCGAC
	[d]A:	AAAAGTGATGACCATGATGCCTTTCT
	[e]D:	GCTAACCAATGCCACTATCCCCCT
YAL056W:	US:	GATTCATTGGCAGGTCCATTGTCGCATTACTAAATCATAGGCATGGATGTCCACGAGCTCTCT
	DS:	ATATTTCTTTTACATTGCTCCGTCTAGATATCAATCCGCCAATCAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTTAAGATGTGTGAACTGCGTCCGTACGCTGCAGGTCGAC
	A:	AGTAATCATGCAACAAGAAAAACCCG
	D:	TATTTCGCTGGTATTGCGAAGGA
YAL053W:	US:	TTCATTACGATTATATTGACGTGATAAAAAGATTATATAGCCATGGATGTCCACGAGCTCTCT
	DS:	AATATTTATATAATAAGTTGTTACATGTGAGTATATATTGGATTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTTAGAGCAGACGCATTGACAGCGTACGCTGCAGGTCGAC
	A:	TCATTAATGCCTTCACGGGAATACGTA
	D:	GGCAATTGCGTACATTTAAGGAAGCA
YAL051W:	US:	TGTGTTAGAAGCAGGTTAATAATAGATTAGGTTGCGTAAGTCATGGATGTCCACGAGCTCTCT
	DS:	ATTATCGGTCAGTTTCATAAATTCGTGGCTATATAATATTAATTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTGATGACAGTACAGTTGACAGCGTACGCTGCAGGTCGAC
	A:	ACCGCTCAAAAAGGTGTATTATCTCC
	D:	CGTTAGACCTCTTCAGTGGTGTGGCT
YAL038W:	US:	TATTTACAAGACACCAATCAAAACAAATAAAACATCATCACAATGGATGTCCACGAGCTCTCT
	DS:	AATTCAAAAAAATAATATCTTCATTCAATCATGATTCTTTTTTTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTATTGAGAGTCAGTGCCAGATCGTACGCTGCAGGTCGAC
	A:	GCCATCAAAACGATATTCGTTGGC
	D:	TCTCGTTTCCAAAAATGCAACACC
YAL037W:	US:	AGTGCCAATAATTACGCAAAAAGCAAAGGAAATAACACTGCTATGGATGTCCACGAGCTCTCT
	DS:	AACAGTCGTAATAATTAAAAATGTTATGCTACCGAGGGCTCTTCAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTAGCGCGACGTTGATCCAGATCGTACGCTGCAGGTCGAC
	A:	TTCGCGTTTCAAGATTTCAAAGGAT
	D:	TTGTATTCGAAAAGGGGGTATCGG
YAL034C:	US:	ATGTTCACTAAAACTGTTGGTACATCCGTGCACCTCCGCTAGATGGATGTCCACGAGCTCTCT
	DS:	GATAAAAGAAAACCATGTTTAAAATGCATACCACCATGTGTATTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTATGAGCGTCCTACGACTGCTCGTACGCTGCAGGTCGAC
	A:	GCTCTGAGTATGCCCAATGCTTCTT
	D:	GAGGACGCAAGCAGGGGATATAGT
YAL031C:	US:	GAAGCTGAAGGTTGAGGATTTAGACTAGTTTTATATTACAACATGGATGTCCACGAGCTCTCT
	DS:	TTTTTCCAGGGACATAAAGAGTTGTTTTTATAAGGTGCGGAGTTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTAGGCGCGTGCTATATCTCTTCGTACGCTGCAGGTCGAC
	A:	TTTGGCGCAAATTTTCAATGTTTAAC
	D:	TGGGGAAAAATCAAATTCGTGAAGAA
YAL030W:	US:	GAAGCTGAAGGTTGAGGATTTAGACTAGTTTTATATTACAACATGGATGTCCACGAGCTCTCT
	DS:	TTTTTCCAGGGACATAAAGAGTTGTTTTTATAAGGTGCGGAGTTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTAGGCGCGTGCTATATCTCTTCGTACGCTGCAGGTCGAC
	A:	TTTGGCGCAAATTTTCAATGTTTAAC
	D:	TGGGGAAAAATCAAATTCGTGAAGAA
YAL028W:	US:	TGATGGCAGTTTTTTTACGTAGTCCAGTAGTTGTCCAGGTACAATGGATGTCCACGAGCTCTCT
	DS:	AAAAAATCACAGGATCATTTTTTTGATATACAAATACTATTTTTTTAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTGGTCTGCTCAAATACACCAACGTACGCTGCAGGTCGAC
	A:	AGGGGAAGGTCAAAAAGCTATTGTT
	D:	TGGTGCCATTTTTCACGGTTACTG

Table 2 (ctd.) Oligonucleotides used for ORF replacement.

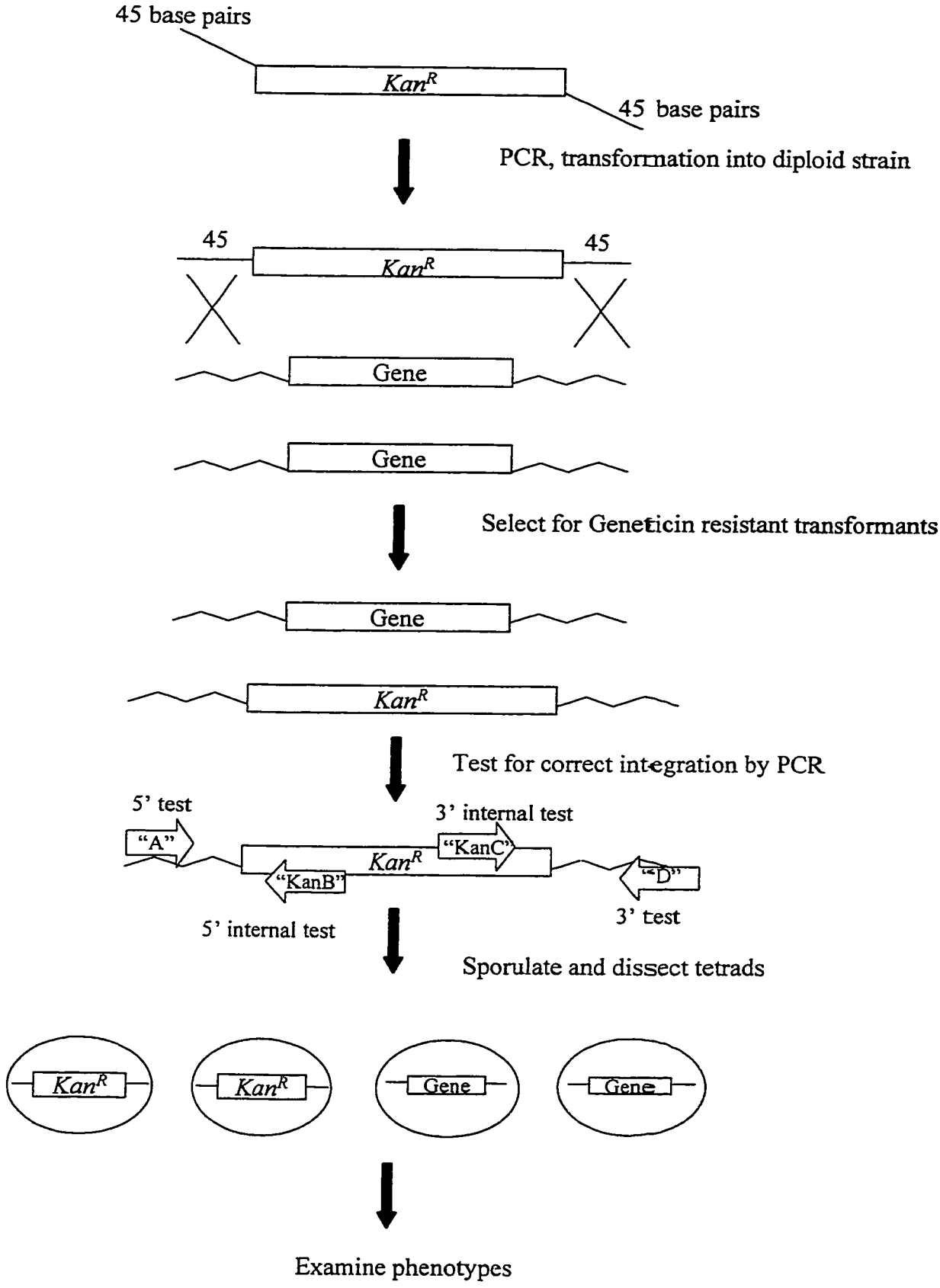
YAL023C:	US:	CATCGCTACTGTCATTCTGTTTCGAACCGTGGTCCAGCAATCATGGATGTCCACGAGCTCTCT
	DS:	ATTCTTGTTTTTTACTTATGTACTTACTACGATGATCTTAATTCAATCGATGAATTCGAGCTCG
	UT:	GATGTCCACGAGGTCTCTCTGCCATGCGAACAAGAGCACGTACGCTGCAGGTCGAC
	A:	ACATATTATTGGACACGTCGCCC
	D:	CAAAGATGAGTAAGGAGAACAGCCTG
YOR375C	US:	ATTATTCTAATATAACAGTTAGGAGACCAAAAAGAAAAAGAAATGCGTACGCTGCAGGTCGAC
	DS:	GAACTTTTTATGAACTTTCCTCTTTTCTTTCTTTTAGACTATTTAATCGATGAATTCGAGCTCG
	A:	GAAGAAGAAAGCGTAAGAAGTAGCA
	D:	AATTGGACGAGTAAGGTCATCAATA
YOR371C	US:	CGGCTACTTTAAGGCTTTCCGTACCAATTCTTCTACATAAGAATGCGTACGCTGCAGGTCGAC
	DS:	TTGTCACATAAGGATTATCCGAACCCCGCCCCGCGGAATTAATTAATCGATGAATTCGAGCTCG
	A:	CATATTGTTCGTCAGGATGTTCATA
	D:	TATATCCGAAGGCAGTATCCTAGTG
Y0R365C	US:	AATATTTAACCTCTCTGATTTATAGATGTAAAGAGAATAAGTATGCGTACGCTGCAGGTCGAC
	DS:	AAGGCAAATCCTACCGGACGTATTAGTCACGATTTTATTTATTCAATCGATGAATTCGAGCTCG
	A:	AGTGGTAATCCATGCTGTATTTTGT
	D:	TTTTTACCTACTCTATTCATTGCCG
Y0R363C	US:	ATCACTTTCTGATCTTGCTGAACAAAATTTCTTCGATAAAAGATGCGTACGCTGCAGGTCGAC
	DS:	TGAAGTGGCCGTATACATCTCAGGAAAACTGGAAAAGACTTCCTAATCGATGAATTCGAGCTCG
	A:	AAAAGCCGCAGATCAATAAGATT
	D:	TATTGTAAATGGAATAAACTGGCGT
Y0R347C	US:	ATCCTCTACGTCCATTGTAAGATTACAACAAAAGCACTATCGATGCGTACGCTGCAGGTCGAC
	DS:	TTAAATAAAATTAAGTAAAAAAAAATAAGGACTTTAATTTTTACTAATCGATGAATTCGAGCTCG
	A:	AAAGTCAATGAAAAGTTCTCCACAC
	D:	AACGTGGGAAAGAATTTTACTCAAT
YOR342C	US:	AATACAAGCTTGTTTCCTGCATTTTTCGATAGAAAAGAAAAAATGCGTACGCTGCAGGTCGAC
	DS:	TCGTTAGGTATGGACATTGATTTGGCCTTCTTCTTCGTGAAACTAATCGATGAATTCGAGCTCG
	A:	GGTCTTTGCTTGTGTAGTTTTTCAT
	D:	TACAAAGGCTATATTTGAAGGAACG
YOR 338W	US:	ACTAATAATATAGATTAGAAAAGAGAATTTTTTTCGACTGCGATGCGTACGCTGCAGGTCGAC
	DS:	CGATGAAACGATACGCTATCCCAAAGGCGTACCGCCTCTTTCTTAATCGATGAATTCGAGCTCG
	A:	TATAAACTACACGCACGATTTTTCA
	D:	ATTGATAACGTAACAACCGAAGAAA
YOR334W	US:	TATACGCAATCGTAGTGAAAGTGATTTTCAATCAAGCATCTCATGCGTACGCTGCAGGTCGAC
	DS:	TTATGCTAATAATATTGTATAGTTTATCATGATAATCGCCTTTCAATCGATGAATTCGAGCTCG
	A:	TGGAGGAAAGATTGAAATTGTTAAG
	D:	GCTATAAAGGATGCAGTTGATGATT
YOR327C	US:	ACATTCGAAACACTTCCAAATACAAAATAAGAACGCGCAACGATGCGTACGCTGCAGGTCGAC
	DS:	CTGTATATATTTTTTAGAATTAGCATCGGCAACCGATGAGCGTTAATCGATGAATTCGAGCTCG
	A:	AGTGATCTTGGTCACATGATATACG
	D:	TAAGAGAATAGTTGACCTTGTTGCC
YOR326W	US:	TGTACATATAGGACATAAAAACAGCAGATATTACAGCGTATAATGCGTACGCTGCAGGTCGAC
	DS:	TTTTAGCATTCATGTACAATTTTGTTTCTCGCGCCATCAGTTTTAATCGATGAATTCGAGCTCG
	A:	CAATTGCACCTAGTGACACATTTAC
	D:	GTGGTGGAAAGATCTAAAAATGAGA

Table 2 (ctd.) Oligonucleotides used for ORF replacement.

YOR324C	US: TTTTTTTTTTGGTTTAAGTTATAGTTGATACATTAAGTGAAAATGCGTACGCTGCAGGTCGAC
	DS: TGTACGCGTACCATTTTTGTAAAGTGACCTTGTTTTGACTTTTTAATCGATGAATTCGAGCTCG
	A: GGTAATAGTTTTTGCCCTACGTTTT
	D: TGTGCTCCTTTTTGTTCTTACTCTT
YOR321W	US: ATATTTGTGTGCTCCAAATACCGGCAAATAAAATAACAGACAATGCGTACGCTGCAGGTCGAC
	DS: TTTTTTGTGTGGTGAAGAAAAGTTTTTGATCATTCTATATATCTAATCGATGAATTCGAGCTCG
	A: ATCATCCTAGAGGCATTACAGCTAA
	D: CACCACATAGCACAAGAGATATCAG
YPL221w	A: CGCTCGAGAACTAACTTAACCATAC
	D: AAAGAAAAAGACTGCTAAAATCCGT
YGL139w	A: GGACGCTATCAAATTAACTCAAAAA
	D: TTTTTAATTTGTTGTTACACGAGGA
YAL053w-URA3	US: TGTTTCATTACGATTATATTGACGTGATAAAAAGATTATATAGCCC AATGATGGGTAACAAGAGC
YAL053w-URA3	DS: TTAAATATTTATATAATAAGTTGTTACATGTGAGTATATATTGGAG CGGT ATTTCACACCGCATA
[f]KanB: CTG CAG CGA GGA GCC GTA AT	
[g]KanC: TGA TTT TGA TGA CGA GCG TAA T	

[a] US: (upstream) 5' primer; [b]DS: (downstream) 3' primer; [c]UT: (uptag) 5' barcode containing primer; [d]A: 5' confirmation primer; [e]D: 3' confirmation primer; [f] KanB and [g]KanC: 3' and 5' universal confirmation primers within the Kanamycin cassette.

Fig. 3. Gene replacement was carried out by transformation of a Kan^R cassette into a diploid strain, followed by homologous integration at regions flanking the ORF of interest. Transformants were selected by Geneticin resistance. The correct integration was verified by PCR, and sporulated to observe potential phenotypes in haploid deletion mutants.


45 base pairs
Kan^R
45 base pairs
PCR, transformation into diploid strain
45
Kan^R
45
Gene
Gene
Select for Geneticin resistant transformants
Gene
Kan^R
Test for correct integration by PCR
5' test
"A"
3' internal test
"KanC"
Kan^R
"KanB"
"D"
5' internal test
3' test
Sporulate and dissect tetrads
Kan^R
Kan^R
Gene
Gene
Examine phenotypes

round of PCR served to add a 20 nucleotide gene specific barcode to the cassette while the second round added regions homologous to the flanking areas of the ORF (in order to allow integration of the cassette by homologous recombination). These primer sequences are available through the Yeast Deletion Project (http://www_sequence.stanford.edu/group/yeast_deletion_project/deletions3.html). The remaining deletion mutants were created using only 5' "upstream" and 3' "downstream" primers. "Upstream" primers all contain 21 bases of the 5' region of the KanMX4 cassette (5'ATGCGTACGCTGCAGGTCGAC 3') preceded by 42 bases of the region immediately upstream the ORF start codon. "Downstream" primers contain 20 bases of the 3' region of the KanMX4 cassette (5'AATCGATGAATTCGAGCTCG 3') followed by 44 bases of the sequence immediately downstream the ORF stop codon. Proper integration by homologous recombination of the KanMX4 cassette was verified by PCR using primers that hybridize within the cassette (universal primers "KanB" and "KanC"), which produced bands of predictable size when amplified with primers hybridizing to flanking sequences upstream and downstream of each ORF ("A" and "D").

All double and some triple mutants were created by mating haploid single or double disruptants, sporulating the diploids, and dissecting the resulting tetrads using an MSM Singer micromanipulator. Yeast DNA was obtained by the protocol described in Ushinsky et al. (1997). The ORF replacements were verified by PCR by amplifying yeast DNA using at least two of the three oligonucleotide pairs: A-KanB, KanC-D, A-D. When possible, two independent isolates of single and

double deletion mutants were created and characterized to be confident of the phenotypic results.

2.1.2 Yeast transformations

In order to obtain two mutants which by tetrad dissection appeared to have lethal phenotypes (*Δyal053w Δygl139w Δypl221w* and *Δyal053w Δygl139w Δypl221w Δyor365c*), diploid double and triple homozygous mutants were transformed with a *URA3* cassette to replace *YAL053w*, using a method similar to the kanamycin cassette replacements. To obtain the *URA3* cassette, *E. coli/ S. cerevisiae* shuttle vector YEp352 (Hill et al. 1993) was primed with two 65-mer oligonucleotides, which contained a 20 base pair region homologous to YEp352 flanking the *URA3* gene (far enough upstream and downstream to contain promoter and termination sequences). Each oligonucleotide contained a 45 base pair targeting sequence on its 5' end homologous to the flanking regions of *YAL053w*. The PCR conditions used were: 1. 94°C for 1 min. 30 sec.; 2. 94°C for 1 min.; 3. 40°C for 1 min.; 4. 72°C for 1 min. 45 sec.; 5. Repeat 2 to 4, 29 times; 6. 72°C for 4 min.

A 1.1 kb DNA fragment was verified on a 1% agarose gel, and quantified visually by comparing with a known quantity of DNA ladder. It was excised from the gel in order to be free of the plasmid template (which conforms *URA3* prototrophy and is easily transformed). The band was purified with a QIAquick Gel Extraction Kit (Qiagen Inc., Canada) and then used to transform yeast strains

5347-48 (*Δygl139w/Δygl139w Δypl221w/Δypl221w* and 5344-46 (*Δygl139w/Δygl139w Δypl221w/Δypl221w Δyor365c/Δyor365c*) by the lithium acetate method (Asubel et al. 1988). Transformants were selected on SD, met+, ura- plates (as the host was met-) and streaked for single colonies. The diploid transformants were verified by PCR for being heterozygous for the *URA3* gene having integrated in the chromosomal DNA at the *YAL053w* site. The transformants were sporulated and then dissected on YEPD plates to observe the haploid phenotypes of the triple and quadruple mutants. Colonies from germinated spores were patched on SD met+, ura- plates to determine if they were uracil auxotrophs.

2.2 Phenotypic tests

Colony purified single and double disruption strains were maintained as patches on YEPD plates [1% (w/v) yeast extract, 2% (w/v) bactopeptone, 2% (w/v) dextrose and 2% (w/v) agar]. Cells from fresh patches were spotted, patched, and streaked for single colonies. For spotting, cells were diluted to an OD_{600} of 0.5, 0.05 and 0.005, and 0.75 µl of each dilution was spotted on various media. The YEPD medium contained no additional supplements for UV and temperature tests. Concentrations of substances (and UV exposure levels) were initially tested on wild types to assess levels of sensitivity. The maximum concentration of test substance (or exposure levels) used was the point where the wild type strains became sensitive. Growth was visually scored as being less than, or equal to wild

type strains. Strains in all tests, except where indicated otherwise were incubated at 30°C until growth was apparent.

For tests involving NaCl, NaCl was added to YEPD to a final concentration of 0.3, 0.6 or 0.9 M prior to autoclaving. For YEPD medium containing cycloheximide, caffeine, or calcofluor white (CFW), all solutions were filter sterilized and added to YEPD after autoclaving and allowing the medium to cool. For cycloheximide tests, a stock concentration of 10 mg/ml in 95% ethanol was added to obtain concentrations of 0.05, 0.10 and 0.25 mg/L cycloheximide. Caffeine was added to obtain concentrations of 5, 10, 15, and 20 mM. A 1% stock of CFW was added for a final concentration of 1.5 mg/ml. UV sensitivity was assessed by exposing cells on YEPD plates to 0, 4000, 7000 or 9000 $\mu J/cm^2$ in a Stratagene 1800 crosslinker and visually scoring for sensitivity after incubation at 30°C. Temperature tests on YEPD involved incubation at 8°C, 30°C and 37°C. Growth at 30°C and 37°C was also tested on SD medium [0.17% (w/v) YNB without amino acids and ammonium sulfate (Difco), 0.5% (w/v) $(NH_4)_2SO_4$, 2% (w/v) dextrose, 2% (w/v) agar, 1 pellet NaOH and required amino acid supplements]. Non-fermentable medium (glycerol plates) contained 1% (w/v) yeast extract, 2% (w/v) bactopeptone, 2% (v/v) glycerol. Nitrogen starvation* tests were performed on SD plates without $(NH_4)_2SO_4$, and glutamate plates were composed of SD with 10 mM glutamate in place of $(NH_4)_2SO_4$. Oleate plates contained 0.67% YNB without amino acids, 0.1% yeast extract, 0.5% potassium phosphate, 0.125% oleic acid in 0.5% Tween 80, 2% agar, and were poured thinly

to better observe halos around colonies. Hydrogen peroxide (H_2O_2) plates contained YEPD with 6 mM H_2O_2, freshly made. For heat shock* tests, cells were replica plated on pre-warmed 54°C minimal medium plates, and kept at this temperature for 20 minutes before being transferred to 30°C. Cell and colony morphology* were examined on SD plates.

2.3 Computer analyses

2.3.1 Sequence alignments

Gene Jockey II Sequence Processing software (Cambridge, U.K.) was used to create dot matrices. For DNA sequences, the matrices were created using a search criteria of 22 matches per 50 nucleotides along the sequence. For amino acid dot matrices, a search criteria of 3 identities out of 6 amino acids was used. FASTA searches (Pearson and Lipman 1988) were carried out from the *Saccharomyces* Genome Database (http://genome-www.stanford.edu) using the translated protein database and the default settings. ClustalX (Thompson et al. 1997) was used for the *YAL031/YOR334(MRS2)* and *YAL037w/YOR342c* pair alignments.

PSORT (http://www.psort.nibb.ac.jp; Nakai and Kanehisa 1992) and PROSITE (http://expasy.nhri.org.tw/tools/scnpsit1.html; Hofmann et al. 1999) were used to predict cellular localization of proteins and signature patterns of amino acid sequences.

---

*Tests performed by Sophia Ushinsky

2.3.2 Cladistic analysis

Parsimony analysis using Paup 3.1.1 (Dave Swofford, Sinauer Associates Inc, Sunderland, Mass.) was performed on duplicated genes in "block 2" which had two or more homologs in other regions of the genome. Sequence alignments made with ClustalX were analyzed on a Power Macintosh computer. Uninformative sites in alignments were excluded and an exhaustive search was used for the analysis. Bootstrap analysis with 100 replicates was performed to evaluate the robustness of clades. The closest related taxa of fungi for which adequate sequence data was available was *S. pombe*. Therefore, homologous sequences of *S. pombe* were used as outgroups.

2.3.3 Percentage identity (% ID)

Percentage identity of amino acids taken from MIPS was determined as follows. Sequences were aligned with Gapped BLAST (v2.0.6) with SEG filtering. Alignments that had an expectation score of less than or equal to 1e-3 were refined by Smith-Waterman alignment algoithms from the USC Sequence alignment package v2.0, with no filtering. Only those alignments that produced at least 20% identity or at least 40% similarity were available.

In order to obtain the remaining percentage ID values for sequences in gene families with lower percentage identities, a rough "higher" estimate was used in this study. Sequences were aligned on Clustal X and the number of identical amino

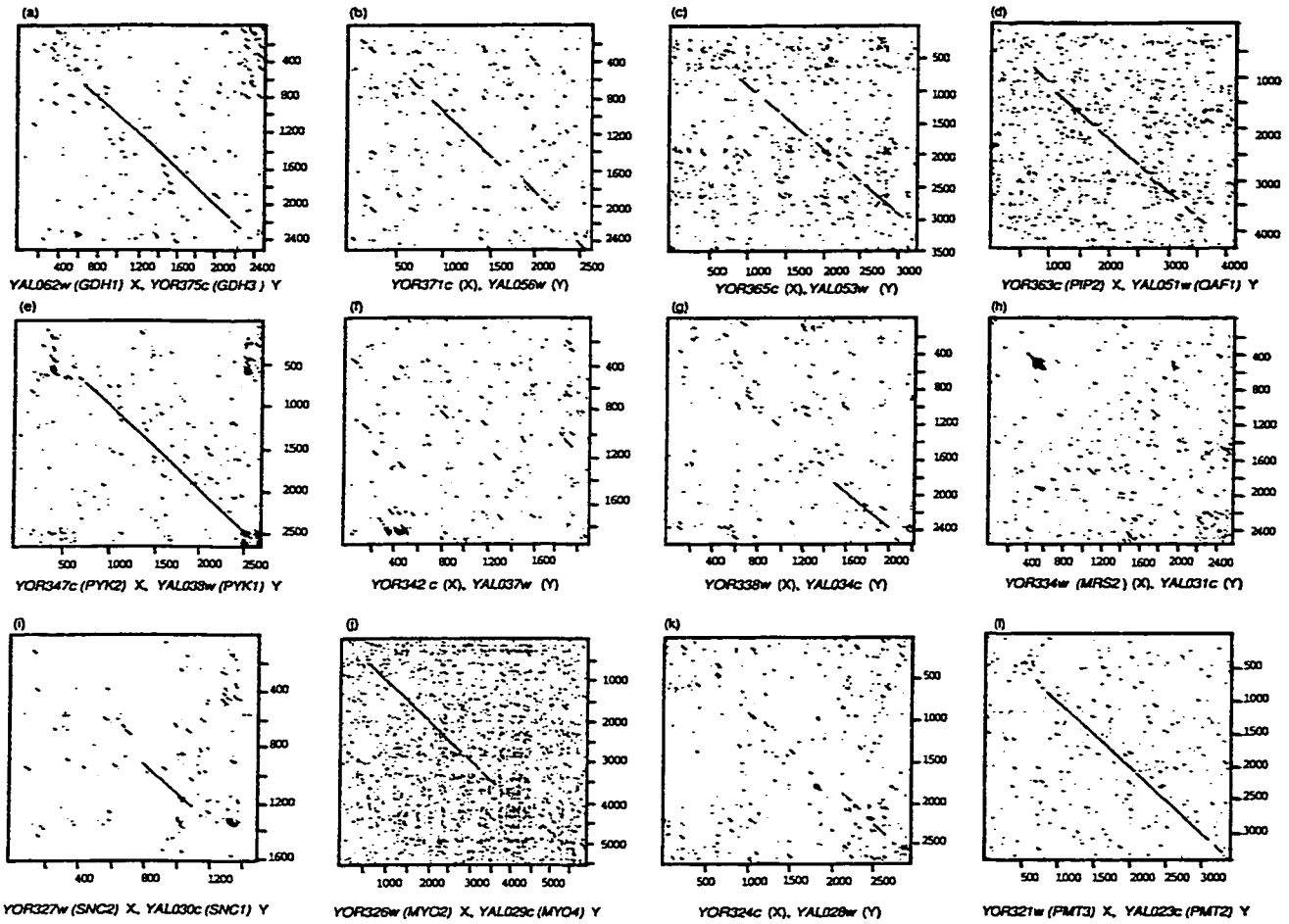
acids was divided by the length of the shorter amino acid sequence. Where indicated, a "lower" estimate was also given (calculated similarly but using the length of the shorter sequence as the divisor).

3. RESULTS AND DISCUSSION

3.1 Sequence homology

Significant similarity was detected in both nucleotide and protein matrix alignments for nine out of the 12 gene pairs aligned [Fig. 4(A),(B)]. However, FASTA detected some similarity between protein sequences coded for by the three gene pairs undetected by the matrix analysis. The gene pair Yor324p/Yal028p (k), show significant similarity, with a FASTA score of 415 (Table 3) and is a recognized member of " block 2" owing to its BlastP score of over 200 (Wolfe and Shields 1997). Yor342p/Yal037p (f) and Yor334p(Mrs2p)/Yal031p (h) are not recognized members of "block 2" despite being surrounded by other members of the block, and have weak FASTA scores of 134 and 96, respectively. However, alignments generated with ClustalX (see Appendix) showed 20 or 25.5% identity for Yor334p(Mrs2p)/Yal031p, depending on whether the number of identical residues was divided by the longer or shorter sequence. Similarly, Yor342p/Yal037p had a lower and higher estimate of 22 and 27%, comparable to

Figure 4. Dot matrix sequence comparisons between paralogous gene pairs (Panel A) and paralogous protein pairs (Panel B) present in the "block 2" region of synteny as identified by Wolfe et al. (1997). All sequence comparisons were performed with the coding strands. (A) DNA matrix comparisons of each syntenic gene pair performed with each ORF sequence plus 500 base pairs of upstream and downstream sequence immediately adjacent to each ORF. (B) Matrix comparisons of the predicted amino acid sequences encoded by each syntenic gene pair.


(a)
YAL062w (GDH1) X, YOR375c (GDH3) Y
(b)
YOR371c (X), YAL056w (Y)
(c)
YOR365c (X), YAL053w (Y)
(d)
YOR363c (PIP2) X, YAL051w (OAF1) Y
(e)
YOR347c (PYK2) X, YAL038w (PYK1) Y
(f)
YOR342 c (X), YAL037w (Y)
(g)
YOR338w (X), YAL034c (Y)
(h)
YOR334w (MRS2) (X), YAL031c (Y)
(i)
YOR327w (SNC2) X, YAL030c (SNC1) Y
(j)
YOR326w (MYO2) X, YAL029c (MYO4) Y
(k)
YOR324c (X), YAL028w (Y)
(l)
YOR321w (PMT3) X, YAL023c (PMT2) Y

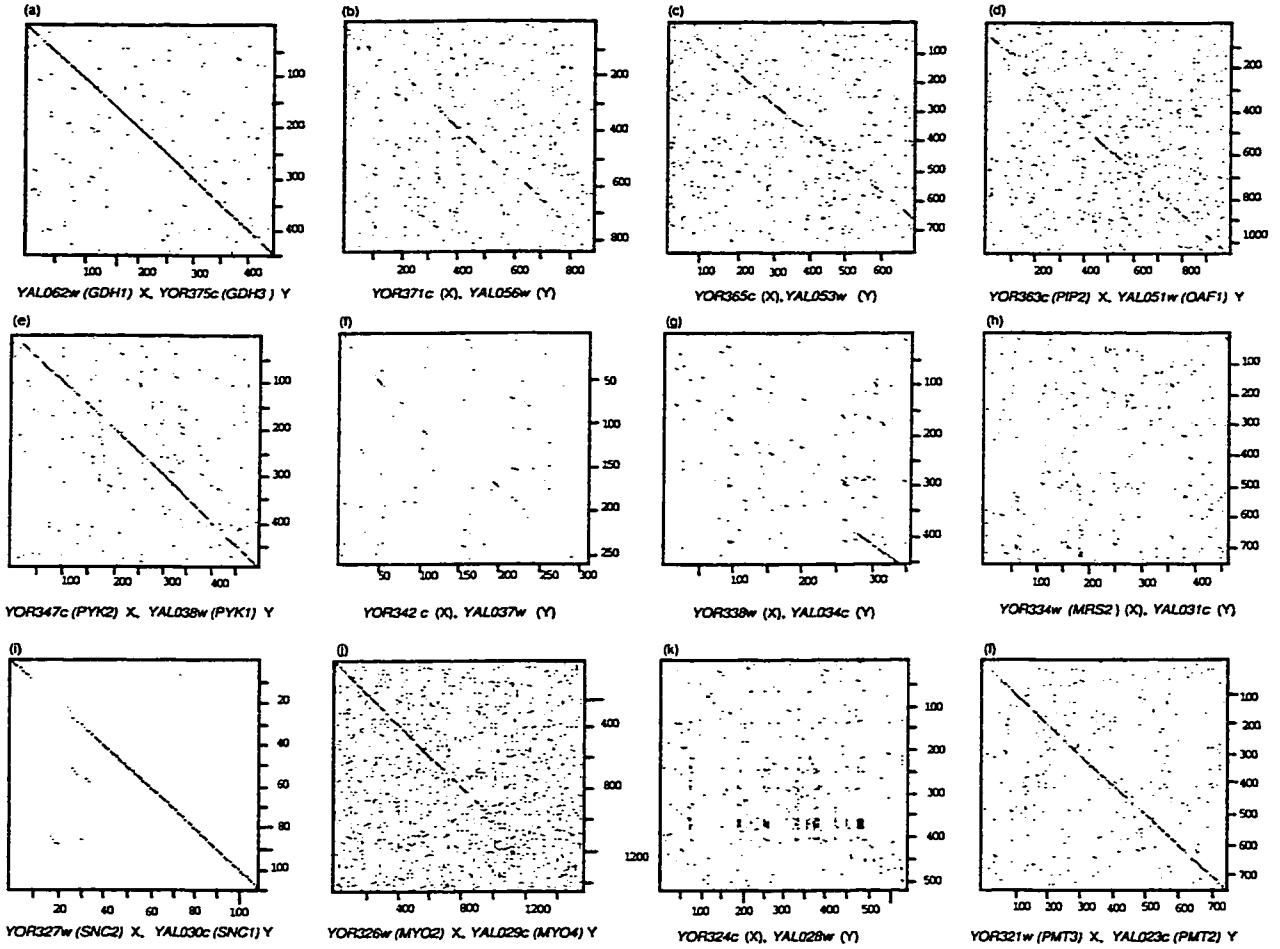

(a)
YAL062w (GDH1) X, YOR375c (GDH3) Y
(b)
YOR371c (X), YAL056w (Y)
(c)
YOR365c (X), YAL053w (Y)
(d)
YOR363c (PIP2) X, YAL051w (OAF1) Y
(e)
YOR347c (PYK2) X, YAL038w (PYK1) Y
(f)
YOR342 c (X), YAL037w (Y)
(g)
YOR338w (X), YAL034c (Y)
(h)
YOR334w (MRS2) (X), YAL031c (Y)
(i)
YOR327w (SNC2) X, YAL030c (SNC1) Y
(j)
YOR326w (MYO2) X, YAL029c (MYO4) Y
(k)
YOR324c (X), YAL028w (Y)
(l)
YOR321w (PMT3) X, YAL023c (PMT2) Y

Table 3. Homology between sequences of gene pairs and FASTA scores.

Gene Pair	Beginning of [1]nt homology relative to START	End of nt homology relative to STOP	Beginning of [2]aa homology relative to START	End of aa homology relative to STOP	[3]FASTA optimized self-score	FASTA optimized score for gene pair
YAL062w/YOR375c	+10	-1	+3	+1	2987	2672
(GDH3/GDH1)	+7	-1	+2	+1	2979	
YAL056w/YOR371c	+513	-219	+259	-301	5698	1084
	+555	-306	+272	-309	6042	
YAL053w/YOR365c	+101	-373	+38	-156	5019	2087
	+167	-64	+60	-52	4508	
YAL051w/YOR363c	+168	-677	+58	-164	6983	2208
(OAF1/PIP2)	+45	-272	+17	-167	6462	
YAL038w/YOR347c	-6	-3	+2	+1	3221	2388
(PYK1/PYK2)	+1	-15	+4	-4	3256	
YAL037w/YOR342c	[4]N.H.	N.H.	N.H.	N.H.	1711	134
	N.H.	N.H.	N.H.	N.H.	2102	
YAL034c/YOR338w	+991	+47	+337	+1	3139	800
	+691	+46	+237	+1	2436	
YAL031c/YOR334w	N.H.	N.H.	N.H.	N.H.	4945	96
(MRS2)	N.H.	N.H.	N.H.	N.H.	3000	
YAL030w/YOR327c	+230	-3	+22	-1	746	565
(SNC1/SNC2)	+113	+1	+21	+1	752	
YAL029c/YOR326w	+1	-1933	+1	-772	9515	4081
MYO4/MYO2	+1	-2330	+1	-665	10086	
YAL028w/YOR324c	+1029	-106	N.H.	N.H.	3444	415
	+1242	-115	N.H.	N.H.	3862	
YAL023c/YOR321w	+167	-9	+62	-6	5229	3476
(PMT2/PMT3)	+146	-12	+55	-7	5174	

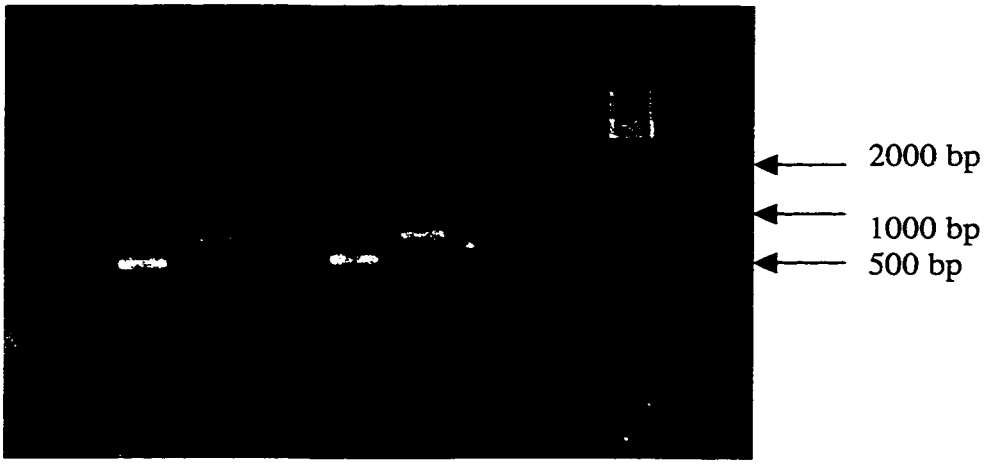
[1]nucleotide; [2]amino acid; [3]aa sequences were used for queries; [4]N.H. denotes no homology was apparent from the dot matrices.

26% as determined by MIPS. Transcription orientation is conserved between *YOR342c* and *YAL037w*, and *YOR334w* and *YAL031c* as in all members of "block 2", and they are located in the same gene order on both chromosomes. For these reasons, we have assumed that these gene pairs are paralogous members of "block 2".

The nucleotide dot matrices [Fig. 4(A)], comparing sequences of the ORFs, and 500 base pairs of the immediately flanking intergenic regions, revealed that sequence similarity was confined to the coding regions of the homologous gene pairs. This suggests that most of the intergenic regions have been saturated with mutations since the duplication that generated "block 2".

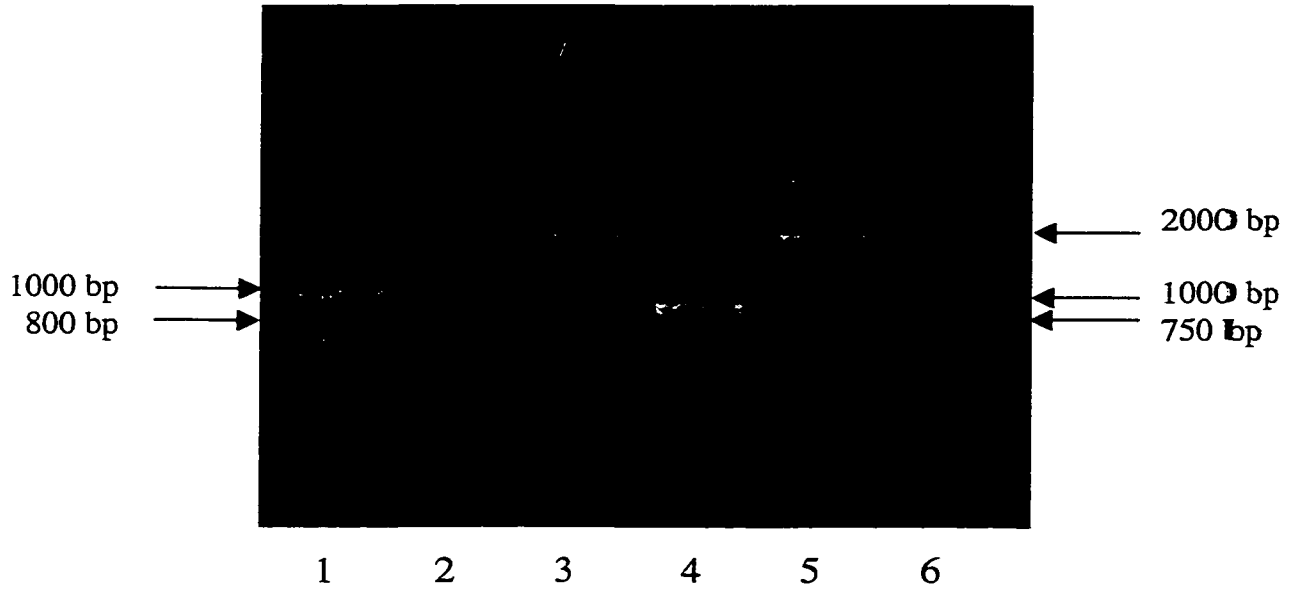
3.2. Strain creation

Strains 5353 [*Δyal031c Δyor334w(Δmrs2)*] and 5304-5 (*ΔpykI Δpyk2*) were created for the current study to complete the collection of mutants which would undergo phenotypic testing (Figs. 5, 6). Since the double mutant 5201-2 (*Δyal053w Δyor365c*) exhibited no phenotype but the genes have two other homologs on other chromosomes (*YPL221w* and *YGL139w*) six mutants were created to test all deletion combinations for phenotypes: 5336 (*Δygl139w Δypl221w*; Fig. 7), 5313 (*Δyal053w Δyor365c Δypl221w*; Fig. 8), 5328-30 (*Δygl139w Δyor365c Δypl221*; Fig.9), 5333-35 (*Δyal053w Δygl139w Δyor365c*;



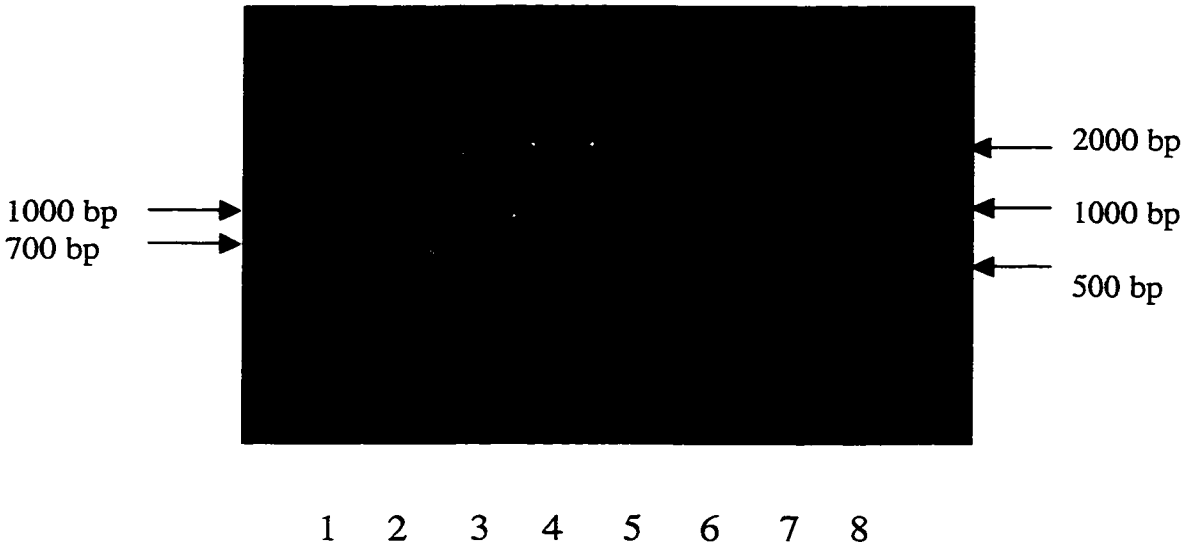
Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	Δ*yor334w*	A-KanB	634
2	Δ*yor334w*	KanC-D	911
3	Δ*yal031c*	A-KanB	525
4	Δ*yal031c*	KanC-D	772
5	1 kb ladder	-	-

Fig. 5. Verification of strain 5353 [Δ*yal031c* Δ*yor334w*(Δ*mrs2*)] by PCR. Below each lane in this and subsequent gels are listed the ORFs verified, the primers used for amplification, and the expected length of PCR product in base pairs.



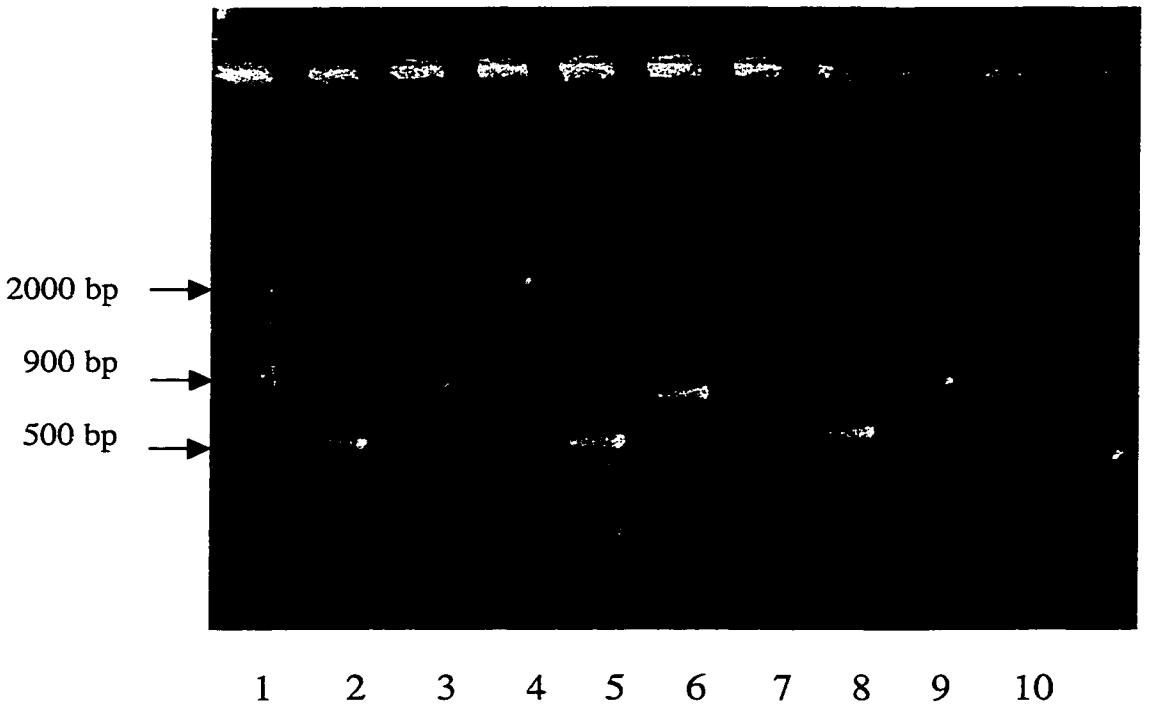
Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	-	-	1 kb ladder
2	*Δpyk1*	KanC-D	813
3	*Δpyk1*	A-D	2008
4	*Δpyk2*	KanC-D	886
5	*Δpyk2*	A-D	2136
6	-	-	1 kb ladder

Fig. 6. Verification of strain 5305 (*Δpyk1 Δpyk2*) by PCR.



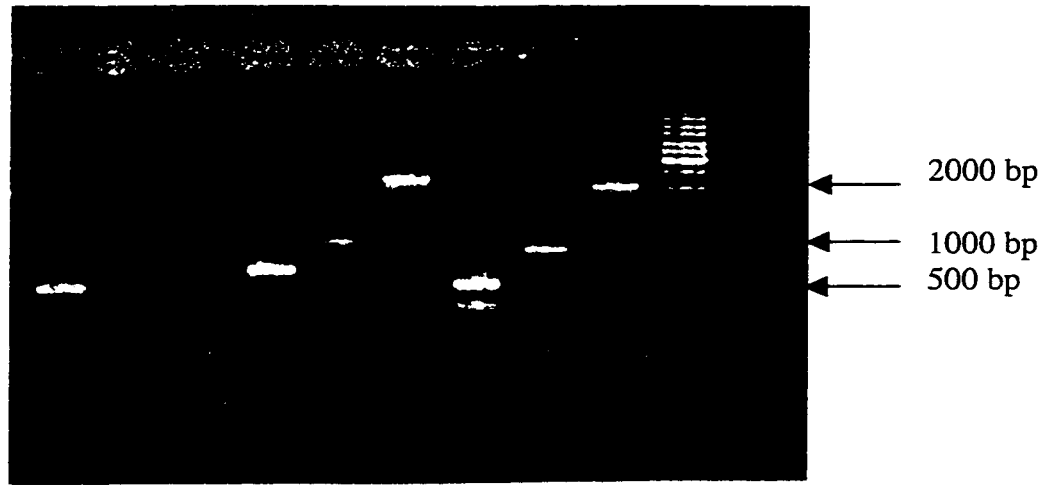
Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	-	-	100 bp ladder
2	Δ*ygl139w*	A-KanB	575
3	Δ*ygl139w*	KanC-D	894
4	Δ*ygl139w*	A-D	2127
5	Δ*ypl221w*	A-KanB	529
6	Δ*ypl221w*	KanC-D	859
7	Δ*ypl221w*	A-D	2046
8	-	-	1 kb ladder

Fig. 7. Verfication of strain 5336 (Δ*ygl139w* Δ*ypl221w*) by PCR.



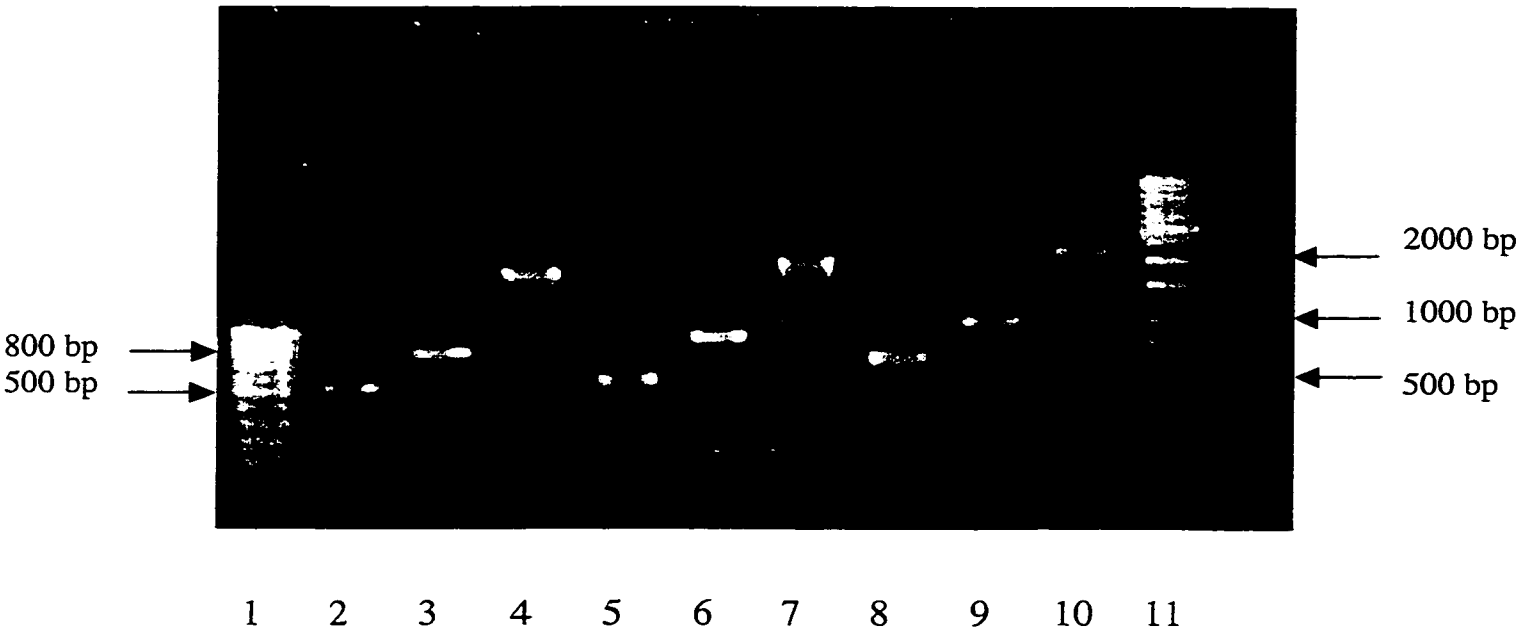
Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	10 kb ladder	-	-
2	*Δyor365c*	A-KanB	561
3	*Δyor365c*	KanC-D	925
4	*Δyor365c*	A-D	2144
5	*Δyal053w*	A-KanB	512
6	*Δyal053w*	KanC-D	792
7	*Δyal053w*	A-D	1951
8	*Δypl221w*	A-KanB	529
9	*Δypl221w*	KanC-D	859
10	*Δypl221w*	A-D	2046

Fig. 8. Verification of strain 5313 (*Δyal053w Δyor365c Δypl221w)* by PCR.



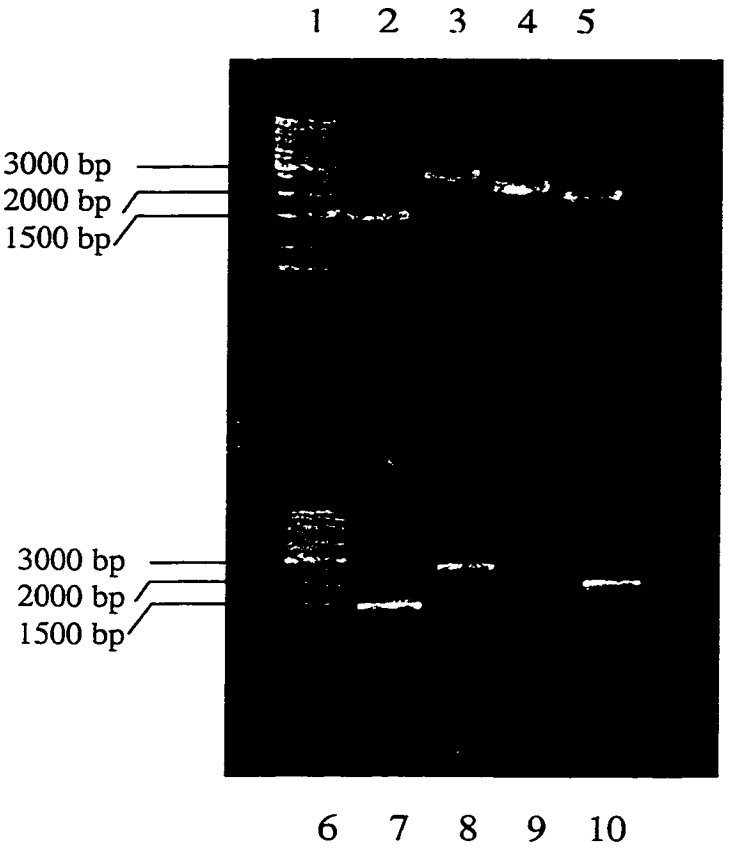
Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	Δ*ypl221w*	A-KanB	529
2	Δ*ypl221w*	KanC-D	859
3	Δ*ypl221w*	A-D	2046
4	Δ*ygl139w*	A-KanB	575
5	Δ*ygl139w*	KanC-D	894
6	Δ*ygl139w*	A-D	2127
7	Δ*yor365c*	A-KanB	561
8	Δ*yor365c*	KanC-D	925
9	Δ*yor365c*	A-D	2144
10	1 kb ladder	-	-

Fig. 9. Verfication of strain 5330 (Δ*ygl139w* Δ*yor365c* Δ*ypl221w*) by PCR.



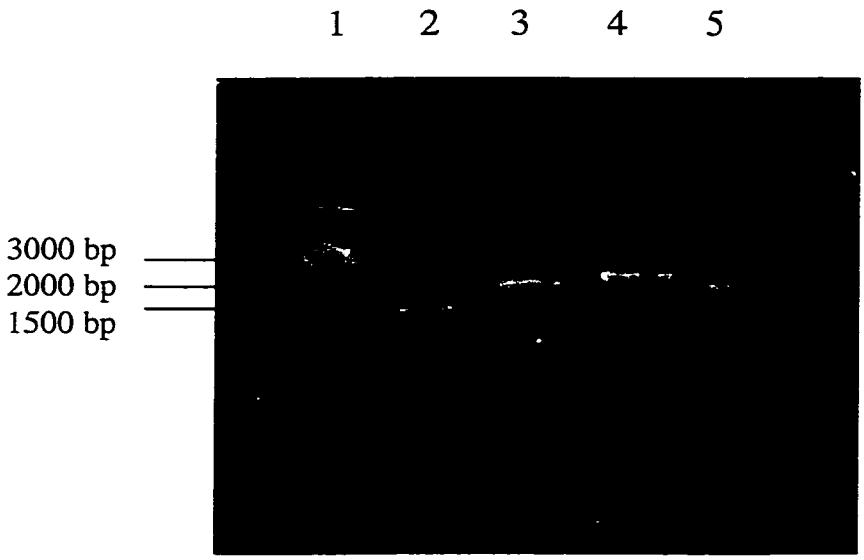
Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	100 bp ladder	-	-
2	Δ*yal053w*	A-KanB	512
3	Δ*yal053w*	KanC-D	792
4	Δ*yal053w*	A-D	1951
5	Δ*yor365c*	A-KanB	561
6	Δ*yor365c*	KanC-D	925
7	Δ*yor365c*	A-D	2144
8	Δ*ygl139w*	A-KanB	575
9	Δ*ygl139w*	KanC-D	894
10	Δ*ygl139w*	A-D	2127
11	1 kb ladder	-	-

Fig. 10. Verification of strain 5335 (Δ*yal053w* Δ*ygl139w* Δ*yor365c*) by PCR.



Lane	Transformed strain	Deletion or wild type confirmed	Primers used	Expected band size (bp)
1	-	-	-	1 kb ladder
2	5349 (Δ*yal053w*/*YAL053w* Δ*ygl139w*/Δ*ygl139w* Δ*ypl221w*/Δ*ypl221*)	Δ*yal053w*	YAL053w A-D	1534 and 2755
3	5349	*YOR365c* (wild type)	YOR365c A-D	2672
4	5349	Δ*ygl139w*	YGL139w A-D	2127
5		Δ*ypl221w*	YPL221w A-D	2046
6	-	-	-	1 kb ladder
7	5350 (same genotype)	Δ*yal053w*	YAL053w A-D	1534 and 2755
8	5350	*YOR365c* (wild type)	YOR365c A-D	2672
9	5350	Δ*ygl139w*	YGL139w A-D	2127
10	5350	Δ*ypl 221w*	YPL221w A-D	2046

Fig. 11. Verification of strains 5349 and 5350 (Δ*yal053w*/*YAL053w* Δ*ygl139w*/Δ*ygl139w* Δ*ypl221w*/Δ*ypl221w*) by PCR. Each strain was derived from an independent transformant. 5349 (above) and 5350 (below) were verified with confirmation primers A-D to have the kanamycin cassettes replacing *YGL139w* and *YPL221w*, but not *YOR365w*. The *URA3* cassette was verified for integration at *YAL053w*.



Lane	Deletion confirmed	Primers used	Expected band size (bp)
1	-	-	1 kb ladder
2	Δ*yal053w*	*YAL053w* A-D	1534 and 2755
3	Δ*yor365c*	*YOR365c* A-D	2144
4	Δ*ygl139w*	*YGL139w* A-D	2127
5	Δ*ypl221w*	*YPL221w* A-D	2046

Fig. 12. Verification of strain 5352 (Δ*yal053w*/*YAL053w* Δ*ygl139w*/Δ*ygl139w* Δ*yor365c*/Δ*yor365c* Δ*ypl221w*/Δ*ypl221w*) by PCR. One transformant was verified with confirmation primers A-D to have the kanamycin cassettes replacing *YGL139w*, *YOR365w* and *YPL221w*. The *URA3* cassette was verified for integration at *YAL053w*.

Fig. 10), 5349-51 (*Δyal053w/YAL053w Δygl139w/Δygl139w Δypl221w/Δypl221w*; Fig. 11), and 5352 (*Δyal053w/YAL053w Δygl139w/Δygl139w Δyor365c/Δyor365c Δypl221w/Δypl221w*; Fig. 12).

3.3.1 Phenotypic analysis

The phenotypic tests chosen, although far from exhaustive, aimed to target a broad range of biological processes. The rationale for each one is briefly listed in Table 4 and is described by Hampsey (1997). Eight out of 12 gene pairs, when singly or doubly disrupted, exhibit phenotypes under the tested conditions. Two gene pairs contain one or both genes that show phenotypes when simultaneously deleted with homologs outside "block 2". Phenotypes were not detected in double mutants for three gene pairs.

In summary (Table 5), out of the 24 singly disrupted null mutants constructed, six (*Δgdh1, Δpyk1, Δmrs2, Δsnc2, Δmyo2,* and*Δ pmt2*) displayed mutant phenotypes. *Δgdh1* grew slowly on minimal medium. *Δpyk1* was unable to grow on fermentable carbon sources. *Δmrs2* was unable to grow on glycerol. *Δsnc2* was sensitive to 0.9 M NaCl and 20 mM caffeine. *Δmyo2* was essential on YEPD at 30°C. *Δpmt2* was sensitive to 0.25 mg/L cycloheximide, 20 mM caffeine and 1.5 mg/ml CFW.

Table 4. Phenotypic tests administered and the metabolic pathways they target.

TEST	METABOLIC PATHWAYS TARGETED
YEPD @ 30°C 37°C 8°C	Fermentation of glucose Heat sensitivity- often defines essential genes Cold sensitivity- assembly of multi-subunit complexes
SD @ 30°C	Amino acid, purine and pyrimidine biosynthesis Defects in transcriptional apparatus
Glycerol/Ethanol	Respiration deficiencies (mitochondrial function)
UV Exposure	Defects in DNA repair genes
Cycloheximide	Protein synthesis inhibitor
NaCl	Defects in cytoskeleton, cell wall, ion channels, pores Vacuolar development Translational fidelity Map kinase signal transduction pathway involving osmosensing
Calcofluor white	Cytoskeletal defects, cell wall defects, secretion
Caffeine	Map kinase pathways, cAMP phosphodiesterase, recombinational repair pathway
Hydrogen peroxide	Genes involved in relief of oxidative stress

Table 5. Results of phenotypic tests

Gene Pair	Ch. I single mutant	Ch. XV single mutant	Double mutant
YAL062w/YOR375c (GDH3/GDH1)	++	SD[1] 30°C(+)	SD 30°C(+)
YAL056w/YOR371c	++	++	Caf[2](s)*, Cyc[3](s)*, H_2O_2[4](s)*, CFW[5](+)*
YAL053w/YOR365c	++	++	++
YAL051w/YOR363c (OAF1/PIP2)	Oleate(h-)	Oleate(h-)	Oleate(h-)
YAL038w/YOR347c (PYK1/PYK2)	YEPD[6](-), SD(-)	++	YEPD(-), SD(-)
YAL037w/YOR342c	++	++	++
YAL034c/YOR338w	++	++	++
YAL031c/YOR334w(MRS2)	++	Gly(-)[7]	Gly(-), H_2O_2(s)*, Caf(s)*, Cyc(s)*
YAL030w/YOR327c (SNC1/SNC2)	++	Caf(s), NaCl[8](s)	inviable**
YAL029c/YOR326w (MYO4/MYO2)	++	inviable	n/a
YAL028w/YOR324c	++	++	++
YAL023c/ YOR321w (PMT2/PMT3)	Caf(s), Cyc(s), CFW(s)	++	inviable**

"++" Wild type growth; "+" Slow growth; "-" No growth; "h-" Lack of halo around cells; "s" Sensitivity to the compound; "*" New synthetic phenotype; "**" Known synthetic phenotype.

[1]Synthetic dextrose (SD); [2]20 mM Caffeine (Caf); [3]0.25 mg/L Cycloheximide (Cyc); [4]6 mM H_2O_2; [5]1.5 mg/mL Calcofluor white (CFW); [6]Rich media (YEPD); [7]Glycerol (Gly); [8]0.9 mM NaCl.

Four double mutants exhibited a synthetic phenotype. Two of these, *Δsnc1 Δsnc2* and *Δpmt2 Δpmt3* showed a synthetic lethal phenotype on YEPD at 30°C. A third double mutant, *Δyal056w Δyor371c* was more sensitive to cycloheximide, caffeine, CFW, and H_2O_2 than either the *Δyal056w* or *Δyor371c* single mutants. The *Δyal031c Δyor334w(Δmrs2)* double mutant strain was more sensitive to cycloheximide, caffeine and H_2O_2 than strains singly disrupted for either of these genes. Double mutant *Δmyo2 Δmyo4* was not created due to the lethality of *Δmyo2* deletions. Triple mutant *Δyal053w Δygl139w Δypl221w* showed a synthetic lethal phenotype on YEPD at 30°C.

From this study, it was concluded that the gene pairs *YAL031c/YOR334w(MRS2), YAL056w/YOR371c*, *SNC1/SNC2*, *PMT2/PMT3* show functional overlap. Two more gene pairs, *YAL053w/YOR365c* and *GDH1/GDH3* show functional overlap between one or both paralogs and additional homologs outside this duplicated region. *YAL053w* appears to overlap with *YGL139w* and *YPL221w* as the triple mutant, *Δyal053w Δygl139w Δypl221w*, has a synthetic lethal phenotype. The triple mutant *Δgdh1 Δgdh3 Δglt1* has a synthetic phenotype of complete glutamate auxotrophy (Avendano et al. 1997). Three gene pairs, *PYK1/PYK2*, *PIP2/OAF1* and *MYO2/MYO4*, were found to encode functionally diverged proteins since mutant phenotypes associated with *PYK1*, *MYO2* and either *PIP2* or *OAF1* deletions are not "masked" by their paralogous gene partner.

Finally, strains with mutations in gene pairs *YAL037w/YOR342c, YAL034c/YOR338w*, and *YAL028w/YOR324c* showed no phenotypes as single or double deletions.

3.3.2 Phenotypic analysis of a paralogous gene family

YAL053w/ YOR365c/ YGL139w/ YPL221w

Strains containing deletions of the *YAL053w* and *YOR365c* genes did not exhibit scorable phenotypes as either single or double mutants. They code for proteins of unknown function that are predicted to localize to the plasma membrane (Table 6). Yal053p and Yor365p share 34-36% identity with Ypl221p and Ygl139p. The possibility that these similar genes perform overlapping functions was explored by creating additional mutants harbouring all possible combinations of triple mutants, as well as a quadruple mutant, for this four member gene family.

First, strain 5313 (*Δyal053w Δyor365c Δypl221w*) was created by crossing 5201-2 (*Δyal053w Δyor365c*) and 5307 (*Δypl221w*), sporulating the resulting diploid and confirming the genotypes of the resulting meiotic products by PCR. Strains 5328-30 (*Δygl139w Δyor365c Δypl221w*) and 5333-35 (*Δyal053w Δygl139w Δyor365c*) were created similarly by crossing strain 5313 (*Δyal053w Δyor365c Δypl221w)* with 5322 (*Δygl139w*). The fourth possible triple mutant *Δyal053w Δygl139w Δypl221w* was not found in any of the 32 spores examined

Table 6. Predicted localization and signature patterns of "block 2" protein encoding genes.

Chromosome I	PSORTII Prediction	Prosite Prediction	Chromosome XV	PSORTII Prediction	Prosite Prediction
YAL062w (GDH3)	73.9% cytoplasmic; 13.0% nuclear	Glu/Leu/Phe/Val dehydrogenase active site (105-118)	*YOR375c (GDH1)*	60.9% cytoplasmic; 13.0% nuclear	Glu/Leu/Phe/Val active site between 104-117
YAL056w	73.9% nuclear; 13.0% cytoskeletal	*	*YOR371c*	52.2% nuclear; 30.4% cytoplasmic	*
YAL053w	47.8% PM; 39.1% ER	ATP/GTP binding site motif A (P-loop) 169-176	*YOR365c*	52.2% PM; 34.8% ER	*
YAL051w (OAF1)	34.8% PM; 17.4% vacuolar; 17.4% nuclear; 17.4% ER	Fungal $Zn_{(2)}$-$Cys_{(6)}$ binuclear cluster domain signature (65-93)	*YOR363c (PIP2)*	47.8% PM; 17.4% nuclear; 13.0% vacuolar	Fungal Zn(2)-Cys(6) binuclear cluster domain signature (24-52); ATP synthase alpha and beta subunits signature (174-183)
YAL038w (PYK1)	52.2% cytoplasmic; 21.7% nuclear	Pyruvate kinase active site signature (235-247)	*YOR347c (PYK2)*	39.1% mitochondrial; 30.4% cytoplasmic	Pyruvate kinase active site signature (237-249)
YAL037w	69.6% nuclear; 13% mitochondrial; 13.0% cytoplasmic	*	*YOR342c*	52.2% nuclear; 26.0% cytoplasmic	*
YAL034c	60,9% mitochondrial; 30.4% nuclear	*	*YOR338w*	65.2% nuclear; 21.7% cytoplasmic	*
YAL031c	39.1% PM; 21.7% nuclear	*	*YOR334w (MRS2)*	47.8% mitochondrial; 39.1% nuclear	*
YAL030w (SNC1)	33.3% vesicles of secretory system; 22.2% PM	Synaptobrevin signature (45-65)	*YOR327c (SNC2)*	33.3% vesicles of secretory system; 22.2% PM	Synaptobrevin signature (45-64)
YAL029c (MYO4)	52.2% cytoplasmic; 34.8% nuclear	ATP/GTP binding site motif A (P-loop) 165-172	*YOR326w (MYO2)*	39.1% cytoplasmic; 39.1% nuclear	ATP/GTP binding site motif A (P-loop) 164-171
YAL028w	26.1% cytoplasmic; 21.7% nuclear	*	*YOR324c*	26.1% cytoplasmic; 17.4% nuclear	*
YAL023c (PMT2)	55.6% ER; 44.4% PM	*	*YOR321w (PMT3)*	65.2% PM; 21.7% ER	*

ER endoplasmic reticulum; PM plasma membrane

*No significant signature patterns.

despite the prediction that equal numbers of each type of triple mutant was expected (1/8 of the spores). In addition, no viable quadruple mutants were found from this cross, although two of the spores were expected to have this genotype.

In a second attempt to improve the chances of finding a quadruple mutant, two of the strains created, 5313 (*Δyal053w Δyor365c Δypl221w*) and 5330 (*Δygl139w Δyor365c Δypl221w*), were mated and sporulated. Since the parent strains each differed by only one ORF replacement, it was expected that each tetratype and non-parental ditype tetrads would include one and two quadruple mutants, respectively. Out of nine tetrads whose genotype was determined by PCR, four were parental ditypes, four were tetratypes, and one was non-parental ditype (Table 7). In all cases of tetratypes and the non-parental ditype, spores with the predicted quadruple mutant genotype did not germinate.

In order to verify that the triple mutant of genotype *Δyal053w Δygl139w Δypl221w* and the quadruple mutant of genotype *Δyal053w Δygl139w Δyor365c Δypl221w* were lethal, the double mutant *Δygl139w Δypl221w* (strain 5336) was created by crossing the two single mutants followed by tetrad analysis. The resulting haploid was mated to a second isolate of the same genotype to make diploid strains 5347-48. Haploid strains 5328-30 (*Δygl139w Δyor365c Δypl221w*) were crossed to each other to create homozygous diploids (strains 5344-46). The two resulting diploid strains were each subjected to gene replacement of the *YAL053w* ORF with the *URA3* cassette (described in section 2.1.2). One out of six

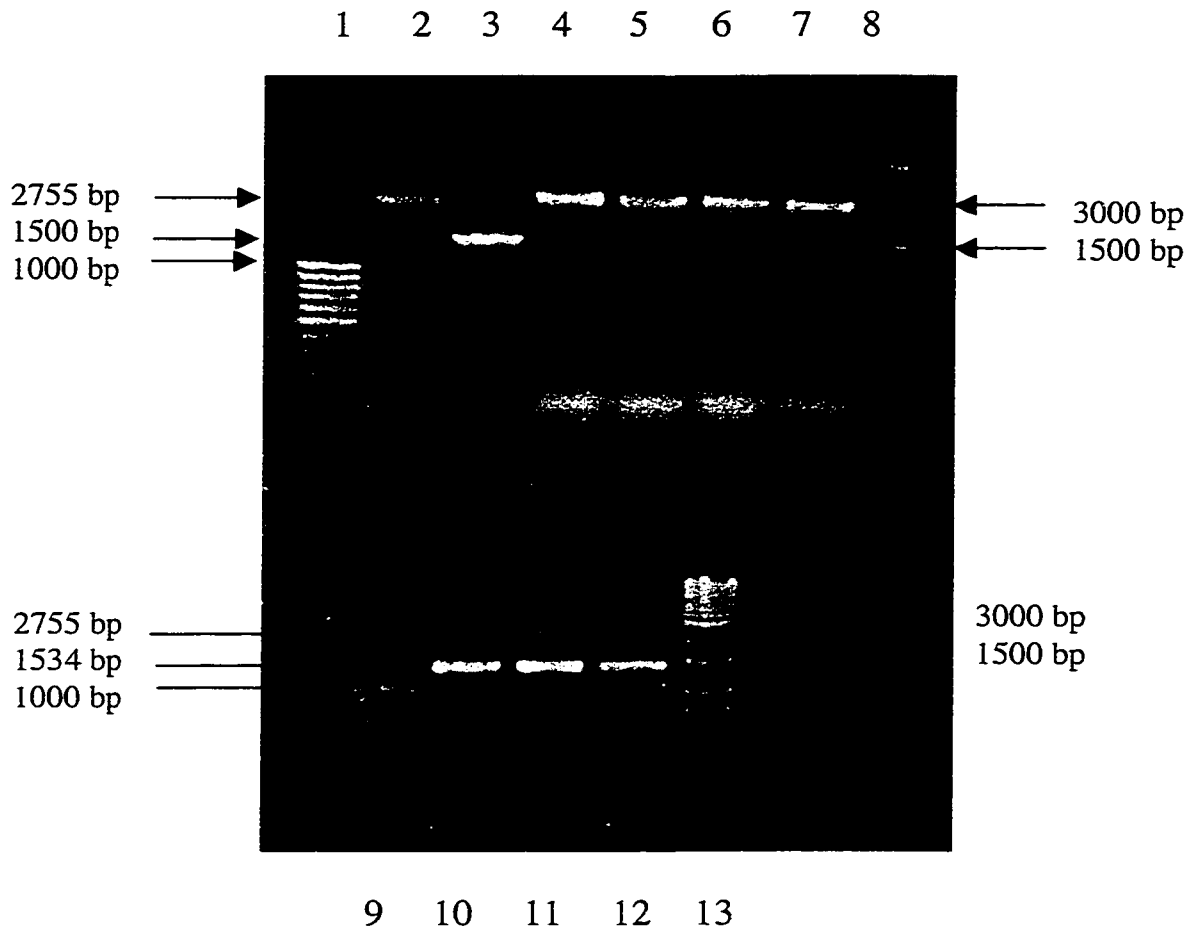
Table 7. Genotypes of spores from nine tetrads obtained by crossing *Δyal053w Δyor365c Δypl221w* to *Δygl139w Δyor365c Δypl22_1w*, as determined by PCR.

GENOTYPE OF SPORES	SPORES ANALYZED
PARENTAL DITYPE: 4 TETRADS	
Δyal053w *Δyor365c Δypl221w*	4
Δyal053w *Δyor365c Δypl221w*	3
Δygl139w *Δyor365c Δypl221w*	4
Δygl139w *Δyor365c Δypl221w*	4
NON-PARENTAL DITYPE: 1 TETRAD	
Δyor365c Δypl221w	1
Δyor365c Δypl221w	1
Δyal053w Δygl139w *Δyor365c Δypl221w* *	0
Δyal053w Δygl139w *Δyor365c Δypl221w* *	0
TETRATYPE: 4 TETRADS	
Δyor365c Δypl221w	4
Δyal053w *Δyor365c Δypl221w*	3
Δygl139w *Δyor365c Δypl221w*	4
Δyal053w Δygl139w *Δyor365c Δypl221w* *	0

*Not viable- genotype implied from surviving spores.

colonies resulting from the transformation of the triple mutant, and all three colonies resulting from the transformation of the double mutant were verified by PCR to have one copy of *YAL053w* replaced by the *URA3* cassette (Fig 13). Since the transformed strains are heterozygous *Δyal053w/YAL053w*, two bands appear when amplifying the DNA with the *YAL053w* A-D primers. One band has 2755 bp and represents the wild type copy (which appears in Fig. 13, but not in Figs. 11 and 12), and the 1534 bp band represents the copy with the *URA3::YAL053w* replacement. Both transformed diploid strains sporulated poorly but about 20 tetrads from each strain were dissected. Dissection of tetrads from strain 5352 (*Δyal053w/YAL053w Δygl139w/Δygl139w Δypl221w/Δypl221w Δyor365w/Δyor365w*) showed 2:2 segregation- two live spores and two dead spores from each tetrad (Fig. 14). Strains 5349-50 (*Δyal053w/YAL053w Δygl139w/Δygl139w Δypl221w/Δypl221w*), were dissected and showed the same segregation pattern following tetrad analysis (Fig. 15). All surviving spores were uracil auxotrophs and therefore did not contain the new gene replacement cassette. These results confirmed that the haploid triple mutant of genotype *Δyal053w Δygl139w Δypl221w* and the haploid quadruple mutant of genotype *Δyal053w Δygl139w Δyor365c Δypl221w* had lethal phenotypes.

Fig. 13. Confirmation of *URA3* replacement of the *YAL053w* ORF by PCR. (Above): Diploid homozygous strain 5344 (*Δygl139w/Δygl139w Δyor365w/Δyor365w Δypl221w/Δypl221w*) was transformed with *URA3* (lane 3) and verified with confirmation primers YAL053-A and D. (Below): Diploid homozygous strain 5347 (*Δygl139w/Δygl139w Δypl221w/Δypl221w*) was transformed similarly (lanes 10, 11, 12) and verified as above.



Lane	Strain undergoing transformation	Primers used	Deletion confirmed	Expected band size (bp)
1	-	-	-	100 bp ladder
2	*5344 (Δygl139w/Δygl139w Δyor365w/Δyor365w Δypl221w/Δypl221w)*	*YAL053w* A & D	*none*	2755
3	*5344*	*YAL053w* A & D	*Δyal053w*	1534 and 2755
4	*5344*	*YAL053w* A & D	*none*	2755
5	*5344*	*YAL053w* A & D	*none*	2755
6	*5344*	*YAL053w* A & D	*none*	2755
7	*5344*	*YAL053w* A & D	*none*	2755
8	-	-	-	1 kb ladder
9	-	-	-	100 bp ladder
10	*5347 (Δygl139w/Δygl139w Δypl221w/Δypl221w)*	*YAL053w* A & D	*Δyal053w*	1534 and 2755
11	5347	*YAL053w* A & D	*Δyal053w*	1534 and 2755
12	5347	*YAL053w* A & D	*Δyal053w*	1534 and 2755
13	-	-	-	1 kb ladder

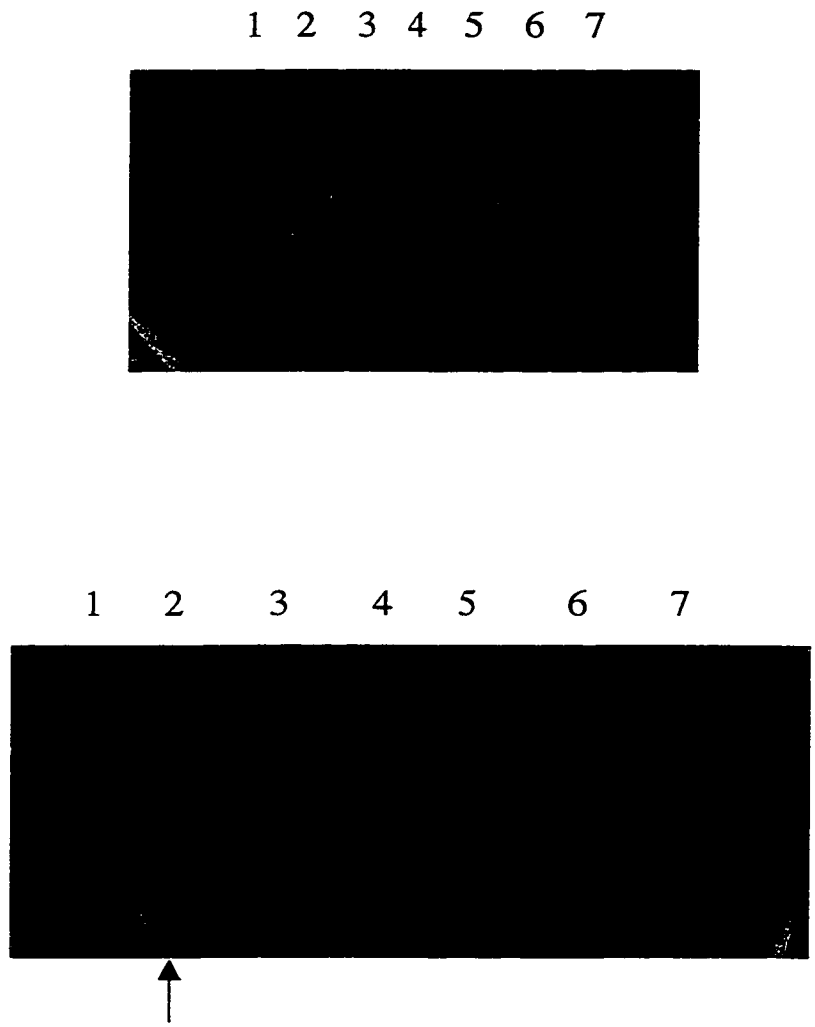


Fig. 14 (Above): Quadruple mutant Δ*yal053w* Δ*ygl139w* Δ*yor365w* Δ*ypl221w* exhibits a synthetic lethal phenotype. The spores from the parent strain 5352 (Δ*yal053w*/*YAL053w* Δ*ygl139w*/Δ*ygl139w* Δ*yor365w*/Δ*yor365w* Δ*ypl221w*/Δ*ypl221w*) heterozygous for Δ*yal053w*, show a 2:2 segregation pattern. (Below): Surviving spores were unable to grow on SD without uracil confirming Δ*YAL053w*::URA3 gene replacement did not occur.

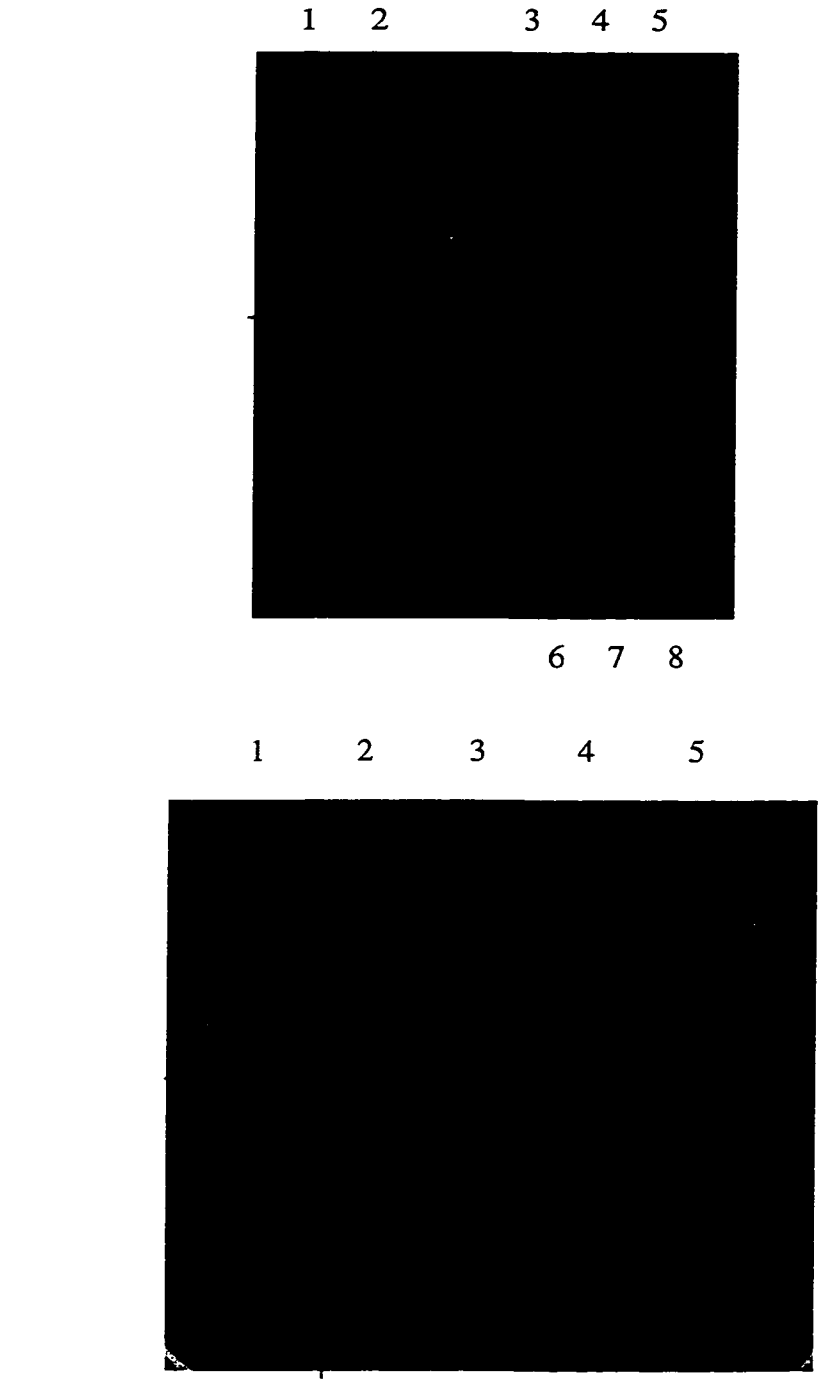


Fig. 15 (Above): Triple mutant Δ*yal053w* Δ*ygl139w* Δ*ypl221w* exhibits a synthetic lethal phenotype. The spores from the parent strain 5349 (Δ*yal053w*/*YAL053w* Δ*ygl139w*/Δ*ygl139w* Δ*ypl221w*/Δ*ypl221w*) heterozygous for Δ*yal053w*, show a 2:2 segregation pattern. (Below): Surviving spores were unable to grow on SD without uracil confirming Δ*YAL053w*::URA3 gene replacement did not occur. One anomalous colony from tetrad 5 exhibiting uracil prototrophy was subsequently confirmed to have been a diploid cell mistaken for a spore.

3.4 The gene pairs in detail

3.4.1 *YAL062w/YOR375c (GDH3/GDH1)*

The *gdh1* single mutant and *gdh1 gdh3* double mutant grew equally slowly on SD, showing an importance for the presence of Gdh1p. *GDH1* mutants typically display a slower than normal growth pattern using ammonium as the sole nitrogen source (Grenson and Hou, 1972; Cooper 1982, Folch et al. 1989; Avendano et al. 1997). These glutamate biosynthesis genes which encode glutamate dehydrogenase $NADP^+$, share 87% identity (Table 8), are similar in function, yet are thought to be utilized in different situations. They use the same co-factor and substrates in the following reaction where the reductive amination of α-ketoglutarate to glutamate occurs (Holzer and Schneider, 1957).

α-ketoglutarate + NH_4 + NADPH ⇨ L-glutamate + H_2O + $NADP^+$

In addition to the two glutamate dehydrogenase- dependent routes, another route for glutamate synthesis exists- the "GS-GOGAT" system. Glutamate is aminated to form glutamine by glutamine synthetase (encoded by *GLN1*). The amide group is then transferred reductively to α-ketoglutarate by glutamate synthase (encoded by *GLT1*) resulting in the conversion of ammonium and α-ketoglutarate to glutamate (Tempest et al., 1970). Although the contributions of these systems under different physiological conditions are not well understood, genetic studies show that they do not contribute equally. For example, Avendano et al. (1997) found that single *gdh1* mutants grown on ammonium as the sole

Table 8. Characteristics of ORFs.

Chromosome I ORF designation	[a]Length	Coordinates	Chromosome XV ORF designation	Length	Coordinates	% Identity	Short Description
YAL062w (GDH3)	457	31568-32941	*YOR375c (GDH1)*	454	1043034- 1041670	87	*GDH3* and *GDH1* : NADP- glutamate dehydrogenase.
YAL056w	847	32960-41803	*YOR371c*	897	1034174- 1031481	35	Function unknown (FUN)
YAL053w	783	45899-48250	*YOR365c*	703	1025564- 1023453	47	FUN-integral membrane protein with 10 potential transmembrane domains
YAL051w (OAF1)	1062	48564-51752	*YOR363c (PIP2)*	1024	1023204-1020214	41	*OAF1* and *PIP2* : Transcription factors required for induction of POX1 on oleate
YAL038w (PYK1)	500	71788-73290	*YOR347c (PYK2)*	506	986453- 9849333	71	*PYK1* : Pyruvate kinase, catalyzes the final step in glycolysis; *PYK2* : glucose repressed isoform.
YAL037w	267	74022-74825	*YOR342c*	319	96716-966657	26	FUN
YAL034c	463	82103-80712	*YOR338w*	363	956889- 957980	42	FUN
YAL031c	760	87033-84751	*YOR334w (MRS2)*	470	944587- 945999	[b]20/25.5	*YAL031c* : FUN; *MRS2* : Protein involved in splicing mitochondrial group II introns and mitochondrial metal ion homeostasis.
YAL030w (SNC1)	117	87288-87754	*YOR327c (SNC2)*	115	931072-930725	77	*SNC1* and *SNC2* : synaptobrevin v-snare homologs present on post-Golgi vesicles.
YAL029c (MYO4)	1471	92272-87857	*YOR326w (MYO2)*	1574	925712-930436	50	*MYO4* and *MYO2* : Myosin class V heavy chain
YAL028w	528	92902-94488	*YOR324c*	602	925031- 923223	28	FUN
YAL023c (PMT2)	758	108550-106274	*YOR321w (PMT3)*	753	916021- 918282	67	*PMT2* and *PMT3* : Transfer mannosyl residues from dolichyl phosphate D-mannose to seryl and threonyl residues in proteins

[a] Length in amino acids; [b]% Identity according to ClustalX alignment (see text for details). All others correspond to values displayed at YPD (http://www.proteome.com/databases).

nitrogen source had a much lower $NADP^+$-GDH specific activity than *gdh3* single mutants. Interestingly, mutants impaired in *GDH1* seemed to compensate for this loss by having an increased GOGAT specific activity, while the GOGAT activity in *GDH3* mutants did not change. Complete glutamate auxotrophy requires *GDH1*, *GDH3*, and *GLT1* to be simultaneously inactivated (Avendano et al. 1997). This information is consistent with the idea that considerable functional overlap exists between *GDH1*, *GDH3* and *GLT1*.

3.4.2 YAL056w/YOR371c

Deletion of one of the members of the *YAL056w/YOR371c* gene pair did not generate a detectable mutant phenotype. However, the double mutant showed significantly increased sensitivity to caffeine, cycloheximide, H_20_2 and CFW (Fig. 16). The *YAL056w/YOR371c* genes therefore provide an example of functional overlap because under certain conditions only one of the two genes is required for wild type levels of resistance to these compounds, but deletion of both genes results in reduced resistance.

3.4.3 YAL051w/YOR363c (OAF1/PIP2)

Previous studies have determined the necessity of both Oaf1 1p and Pip2p for the use of oleate (Karpichev 1997, 1998; Rottensteiner et al. 1997) indicating that they are functionally diverged. As expected, the *Δoaf1* and *Δpip2* single mutants, and the double deletion mutant failed to utilize oleate as a sole carbon

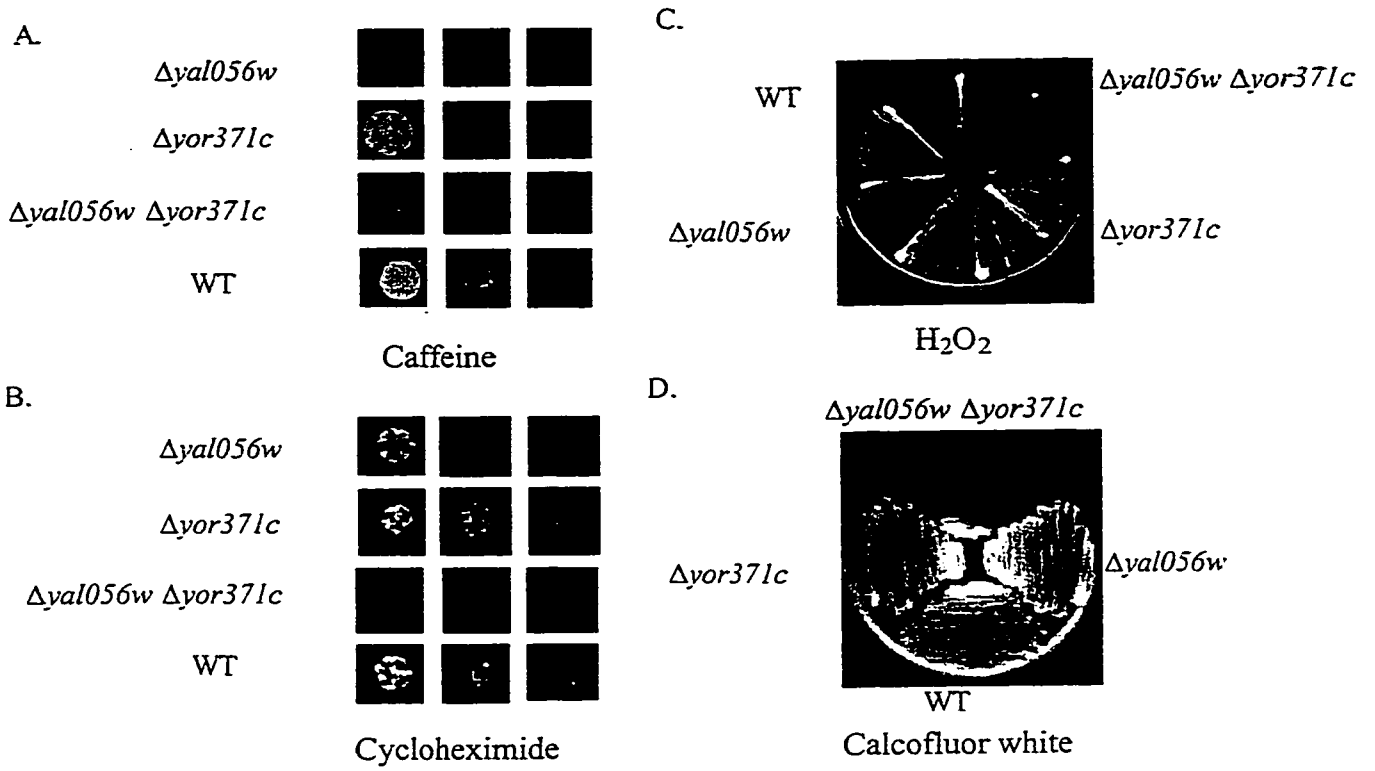


Figure 16. *YAL056w* and *YOR371c* show functional overlap as they appear important for wild type levels of resistance to caffeine, cycloheximide, hydrogen peroxide and calcofluor white. Relative growth of single mutants Δ*yal056w* and Δ*yor371c* and double mutants Δ*yal056w* Δ*yor371c* in the presence of 20 mM caffeine (Panel A), 0.25 mg/L cycloheximide (Panel B), 6 mM hydrogen peroxide (Panel C) and 1.5 mg/ml calcofluor white (Panel D). Spotted cells on this and all figures are serial dilutions corresponding to absorbances 0.5, 0.05, and 0.005 at OD_{600}.

source as evidenced by the lack of clear halos surrounding the colonies (data not shown). Oaf1p and Pip2p are Zn_2Cys_6 transcription factors which positively regulate genes encoding peroxisomal proteins used in lipid metabolism (Luo et al. 1996; Karpichev et al. 1997; Rottensteiner et al. 1996, 1997). In the presence of oleate, both proteins coordinately bind to a UAS, the oleate response element (ORE) as a heterodimer (Karpichev et al. 1997; Rottensteiner et al. 1997). To date, 22 genes, identified as being induced by oleate, require one or both of these transcription factors for activation (Karpichev and Small 1998).

3.4.4 YAL038w/YOR347c (PYK1/PYK2)

The *pyk1* mutants were unable to grow on a fermentable carbon source (glucose) while *pyk2* mutants were able to grow on fermentable as well as non-fermentable media (glycerol), as expected (Boles et al. 1997, 1998). The double mutant, *pyk1 pyk2*, was phenotypically similar to the single mutant *pyk1*.

Wolfe and Shields (1997) suggest that the ancient genome duplication allowed for the evolution of genes dedicated to fermentation of sugars under anaerobic conditions. It seems likely that *PYK1* and *PYK2* constitute a gene pair which once had a common function but diverged, allowing Pyk1p to acquire a specific function for growth on fermentable carbon sources. This is supported by the finding that overexpressed Pyk2p can support growth on glucose of *PYK1* deletion mutants (Boles et al. 1997). This suggests that functional divergence includes changes in the regulation of Pyk1p and Pyk2p.

Pyruvate kinase (Pyk1p) is induced by fructose 1,6-bisphosphate and catalyzes the conversion of phosphoenolpyruvate to pyruvate, the final step in glycolysis. Interestingly, the kinetic properties of Pyk1p suggest it has a low *in vivo* activity during gluconeogenesis (Boles, et al., 1998). On the other hand its isozyme, Pyk2p, with whom it shares 71% identity, is neither induced by fructose 1,6-bisphosphate, nor is it essential for glycolysis (Boles et al. 1997). In fact, Pyk2p is glucose repressed and ethanol induced (Boles et al. 1997). It has been proposed that besides the glycolytic production of pyruvate on fermentable carbon sources, pyruvate can be produced by at least two other pathways during growth on non-fermentable carbon sources such as ethanol or acetate. Boles et al. (1998) believe malic enzyme (encoded by *MAE1*) converts malate to pyruvate by oxidative carboxylation, (which occurs in *pyk1 pyk2* mutants but not *pyk1 pyk2 mae1* mutants) and the other includes the use of (and possible heterodimeric interaction between) pyruvate kinase isozymes, which occurs in *mae1* mutants but not in *mae1 pyk1* or *mae1 pyk2* mutants.

3.4.5 YAL031c/YOR334w (MRS2)

The *Δyor334w(Δmrs2)* single mutant and *Δyal031c Δyor334w(Δmrs2)* double mutant were both unable to grow on glycerol. In addition, the double mutant showed increased sensitivity to caffeine (data not shown), cycloheximide, and H_2O_2, while the single mutants *Δyor334w(Δmrs2)* and *Δyal031c* did not display increased sensitivity to these compounds (Fig. 17). Yal031p is a protein of

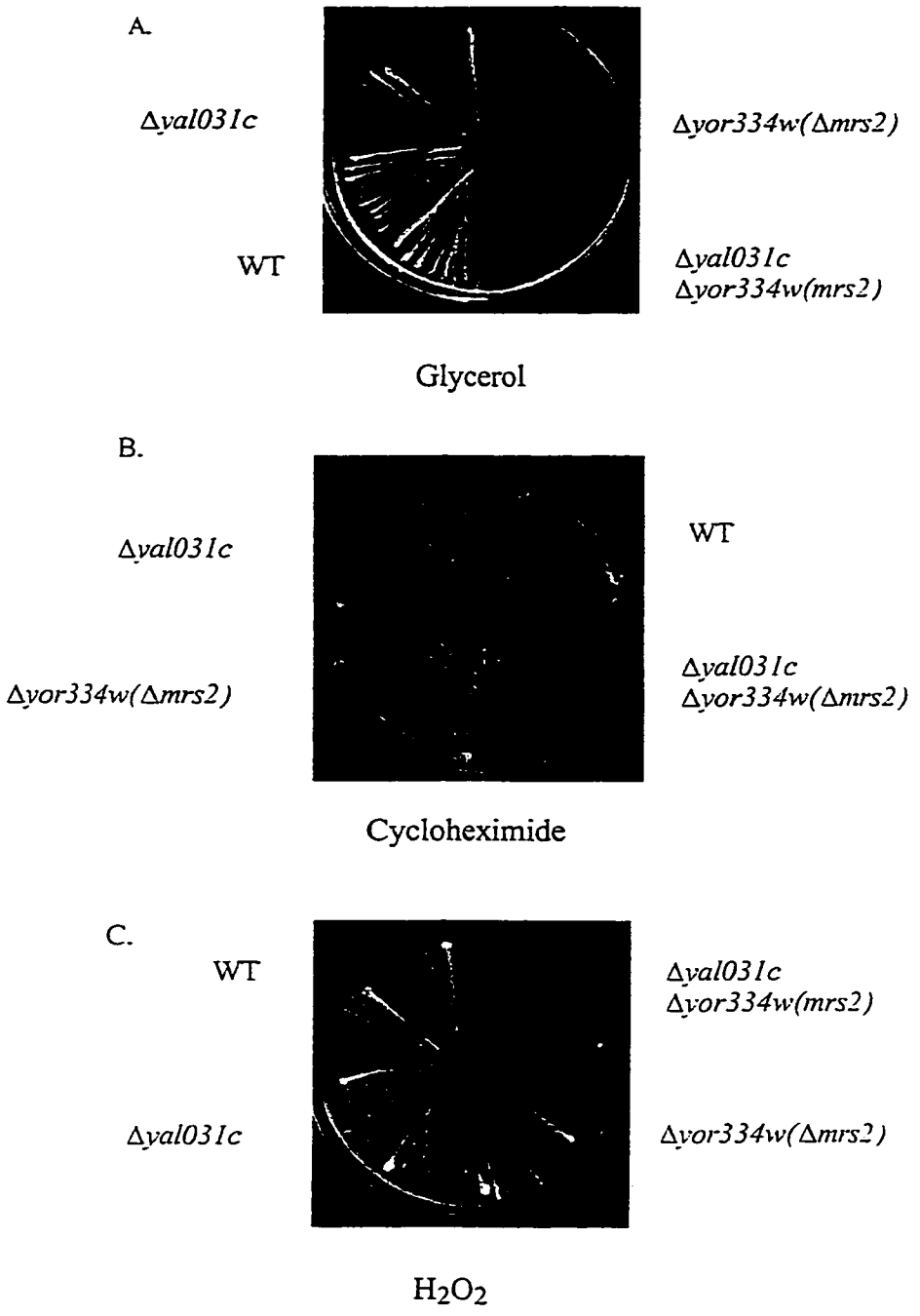


Figure 17. *YOR334w* is required for growth using glycerol as a sole carbon source (Panel A). *YAL031c* and *YOR334w* encode functionally overlapping proteins required for normal levels of cycloheximide and hydrogen peroxide resistance (Panels B and C).

unknown function. *YOR334w* (*MRS2*) encodes a protein which is involved in splicing mitochondrial group II introns (Koll et al. 1987, Wiesenberger et al. 1992, Waldherr et al. 1993) and recent data suggests a role in mitochondrial metal ion homeostasis (Bui et al. 1999). Mrs2p is essential for the formation of a functional respiratory chain, and consequently mutants will not grow on non-fermentable carbon sources such as glycerol (Wiesenberger 1992). It is also 40% identical to the protein Ypl060p of chromosome XVI.

3.4.6 YAL030c/YOR327w (SNC1/SNC2)

Only the single mutant, *Δsnc2*, was sensitive to caffeine and NaCl (Fig. 18). The double mutant, *Δsnc1 Δsnc2*, was inviable on YEPD at 30°C in agreement with Protopopov et al. (1993) who only successfully isolated them on SD below 30°C. They found Snc1p and Snc2p, two homologs of the synaptobrevin/VAMP (vesicle associated membrane protein) family, to be present on post-golgi vesicles in the late secretory pathway and suggested their involvment in targeting and transport of secretory proteins to the plasma membrane. One of the two "V-snares" must be present to interact with the plasma membrane "T-snares" for docking and fusion which is why double accumulate post golgi vesicles and are deficient in bulk secretion.

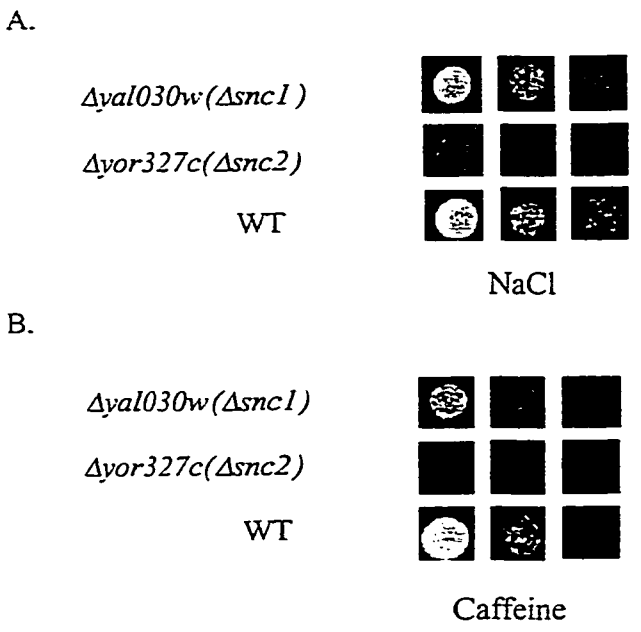


Fig. 18. *SNC2* but not *SNC1* is required for wild type levels of NaCl and caffeine resistance..

3.4.7 YAL029c/ YOR326w (MYO4/MYO2)

MYO2 was found to be essential on YEPD at 30°C (data not shown), in agreement with Johnston et al. (1991) while *Δmyo4* deletion mutants were viable and without phenotypes. *MYO2* and *MYO4* both encode myosin class V heavy chains. *MYO2* is involved in polarized growth and secretion (Johnston et al. 1991; Brockerhoff 1994; Lillie et al. 1994; Govindan et al. 1995) while *MYO4* is required for preferential accumulation of *ASH1* mRNA in daughter cells. Ash1p, an inhibitor of *HO* expression, prevents mating-type switching in daughters (Bobola et al. 1996; Jansen et al. 1996; Long et al. 1997). Both molecules are thought to form dimers with 2 heads, a short coiled-coil domain in the middle and 2 tails (Haarer et al. 1994). Despite their 50% identity and slightly less similarity to other family members (*MYO1*, *MYO3* and *MYO5*), the inability of these genes to functionally replace one another indicates that they code for functionally distinct proteins.

3.4.8 YAL023c/YOR321w (PMT2/PMT3)

The *Δpmt2 Δpmt3* double mutants were inviable on YEPD at 30°C whereas the single mutants were viable, indicating functional overlap between the homologs. *Δpmt2* but not *Δpmt3* single mutants were sensitive to caffeine, cycloheximide and CFW (Fig. 19). *PMT2* and *PMT3* are 2 of a 7 member protein mannosyltransferase family and share the highest identity of any two members (67%). They transfer mannosyl residues from dolichyl phosphate D-mannose to

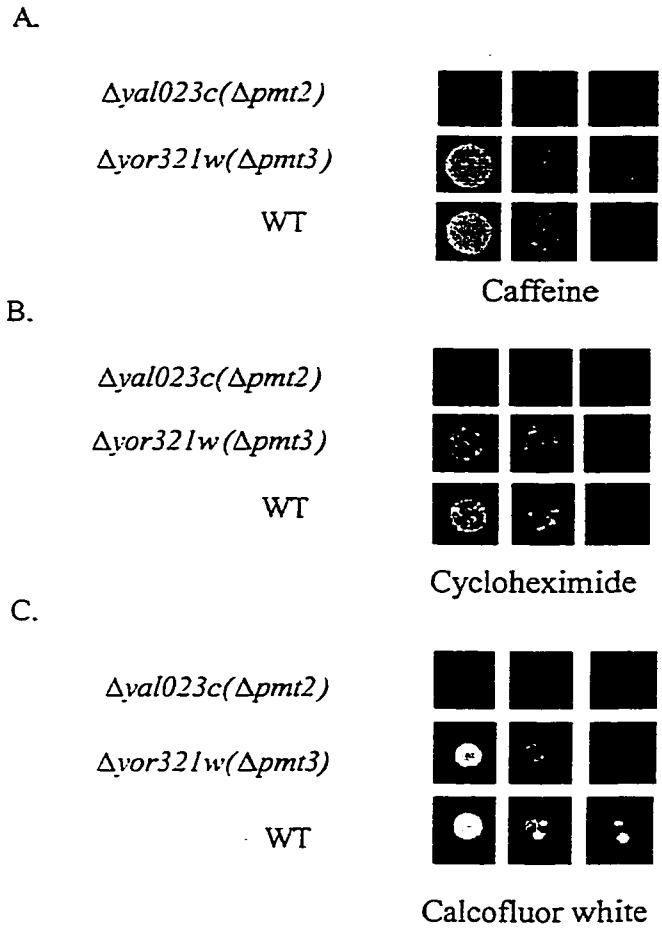


Fig. 19. *Δpmt2* but not *Δpmt3* mutant strains are sensitive to caffeine, cycloheximide and calcofluor white.

seryl and threonyl residues in proteins in the secretory pathway (Immervoll 1995, Lussier et al. 1995). There is evidence that members of this family differ in their specificity towards the protein substrates which they glycosylate. For example, Pmt1p and Pmt2p have been found to glycosylate chitinase, while Pmt4p was found to be involved in o-mannosylation of the plasmalemma-bound protein gp115/Gas1p (Gentzsch and Tanner 1996, Immervoll et al. 1995). Pmt1p and Pmt2p form heterodimers (Gentzsch et al. 1995) and it is suspected that Pmt1p and Pmt3p, as well as Pmt1p and Pmt6p form them (Gentzsch and Tanner, 1996). It would therefore appear that various combinations specify the type of proteins to be glycosylated.

PMT mutants exhibit numerous phenotypes related to cell wall defects. Our results were in agreement with Gentzsch and Tanner (1996) who found a *Δpmt2 Δpmt3* strain to be inviable on YPD at 30°C (unless osmotically stabilized with 1M sorbitol) and sensitive to caffeine and CFW. Caffeine, a diesterase inhibitor is associated with defects in cell wall stability thought to be a result of its effects on the protein kinase C pathway (Levin and Bartlett-Heubusch 1992, Paravicini et al. 1992). Sensitivity to CFW is exacerbated by defects in cell wall biosynthesis and morphogenesis (Ram et al. 1994). McCusker et al. (1988) found most mutations previously isolated as being resistant to low levels of cyloheximide affect permeability. Gentzsch and Tanner (1996) also found *Δpmt2 Δpmt3* mutants to have a higher than normal mannose to glucose ratio in cell walls, abnormal 'clumpy' morphology, grow slowly, and exhibit sensitivity to high temperature

(37°C). PMT mutants, show partial killer toxin resistance because K1 interacts with o-mannosylated protein as well as β-1,6-glucan components (Boone 1990, Lussier et al. 1995). A reduced number of O-linked mannose chains on cell wall mannoproteins provide an inefficient toxin receptor.

3.4.9 YAL053w/YOR365c, YAL037w/YOR342c, YAL034c/YOR338w YAL028w/YOR324c

Four paralogous gene pairs in "block 2" code for proteins of unknown function. They all failed to exhibit phenotypes as double mutants.

However, the combination of the deletion of two homologous genes outside "block 2", *YGL139w* and *YPL221w*, in combination with *YAL053w* resulted in a synthetic lethal phenotype on YEPD at 30°C. Since the fourth homolog, *YOR365c*, was unable to rescue the phenotype, it appears that it does not share gene function with the other three homologs. Since all other triple mutants were viable, it appears that *YOR365c* is not essential for vegetative growth, and that functional overlap exists between the other three homologous genes. Strains 5349, 5350 and 5352 all sporulated poorly, which suggests that *YOR365c* (deleted in strain 5352 but not 5349 or 5350) does not affect sporulation efficiency. It would also seem that *YAL053w* is an important gene for viability because even in single copy it is enough to allow the transformant strains to survive and sporulate. Due to the existence of a poor sporulation phenotype in some mutants, future functional tests

for genes in this family should include comparative measures of sporulation efficiency.

The last three gene pairs bear no sequence similarity to other genes in *S. cerevisiae*. Yal037p is 26% identical to Yor342p while Yal034p is 35% identical to Yor338p and Yal028p shares 32% identity to Yor324p.

The only other information available for these genes is SAGE (serial analysis of gene expression) data. A study profiling the transcriptional program of *S. cerevisiae* during sporulation (Chu et al. 1998) found over 1000 genes whose mRNA levels change during sporulation and therefore may be correlated to this biological process. Included in the induced genes are *YOR338w* and *YOR365c*. Notably, Yor338p does show limited similarity to the AAC11 gene product, which is expressed in *Dictyostelium discoideum* during spore germination (Shaw et al. 1989). Genes repressed during sporulation include *YAL034c* and *YAL053w*. In contrast, *YOR324c* transcript increases in the first hour without significant changes until a decrease follows after 11 hours. A study of mRNA expression during the diauxic shift (deRisi et al. 1997), showed only two of these genes undergoing changes in expression. *YAL034c* increases at the end of the diauxic shift, while *YOR342c* decreases.

3.5 Phylogenetic trees

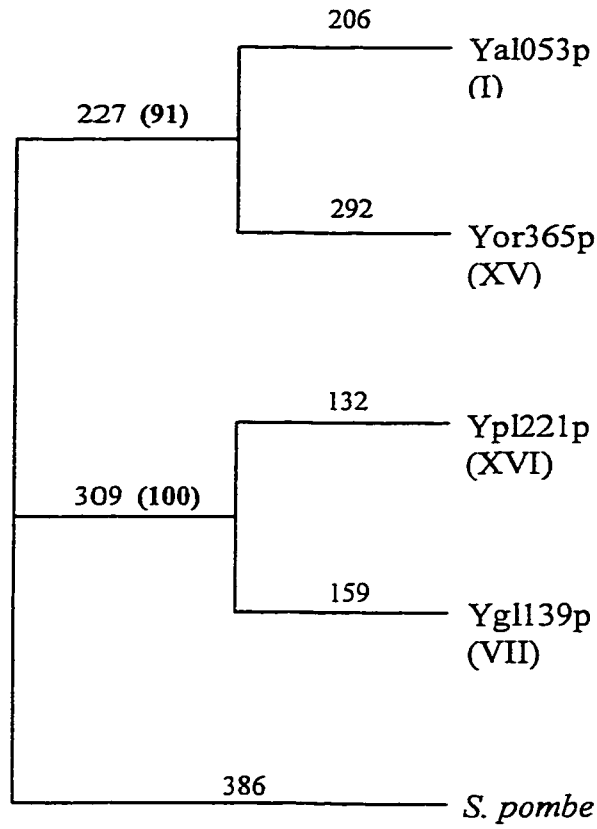
Phylogenetic trees were constructed for four gene pairs that have additional homologs in order to determine the order in which the genes in a family were likely to have evolved.

Both protein pairs Yal053p/Yor365p and Ypl221p/Ygl139p have high identities (47% and 66% respectively- Fig 20B) and are members of "block 2" and "block 32" (Wolfe and Shields 1997). Parsimony analysis of aligned amino acid sequence data using PAUP supports a common ancestor for each clade, both of which may have duplicated at the whole-genome duplication event (Fig. 20A).

The myosin family appears to have originated from two ancestral groups, *MYO1* and an ancestor which gave rise to the remaining myosins. Under the assumption of a whole-genome duplication, the current absence of a homolog for Myo1p suggests a deletion occurred after the duplication. Myo2p/Myo4p and Myo3p/Myo5p share the highest identities of the family (50 and 73% respectively, Fig. 21B) and belong to "blocks 2" and "block 43" (Wolfe and Shields 1997). This is strongly supported by bootstrap data, which clusters each gene pair into a clade 100% of the time (Fig. 21A).

Two ancestors appear to have given rise to the protein mannosyltransferase family- the ancestor of *PMT4* and one other from which the remaining genes originated (Fig. 22A). Since two members, *PMT2/PMT3*, apparently arose from the whole-genome duplication, it is assumed that genes pre-dating these

A.

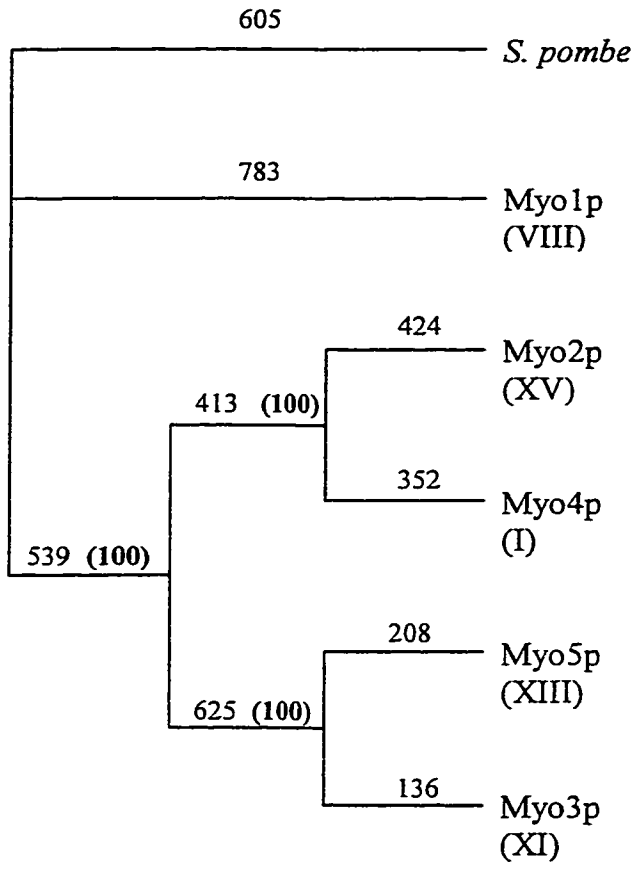


B.

Yal053p	Yor365p	Ygl139p	Ypl221p	
	47 (44.2)	35 (34.4)	36 (34.7)	**Yal053p**
		34 (30.6)	35 (31.4)	**Yor365p**
			66 (65.8)	**Ygl139p**
				Ypl221p

Fig. 20 (a). Phylogenetic tree for unknown proteins Yal053p/Yor365p/Ypl221p/Ygl139p, using an *S. pombe* homolog (accession no. Q09917) as an outgroup. All trees were created with Paup 3.1.1. based on amino acid sequences. Bootstraps were performed 100 times and values are in parentheses. All other numbers indicate branch lengths. Roman numerals indicate the chromosome on which the protein encoding gene is located. (b) Percentage identity of amino acids for Yal053p/Yor365p/Ypl221p/Ygl139p. In all charts values determined by MIPS are left, and percentage identity destimated in this study are in parentheses.

A.

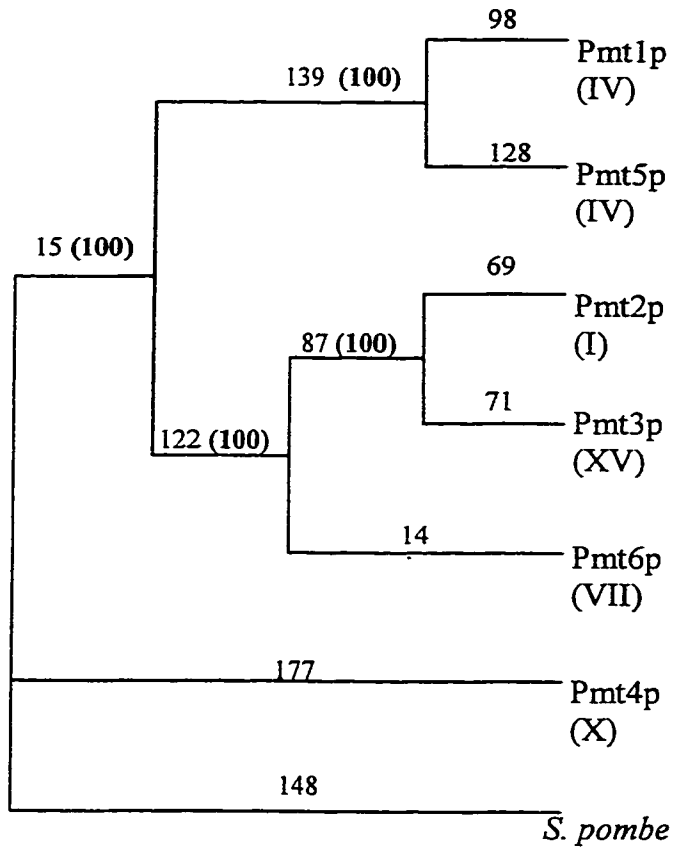


B.

Myo1p	Myo2p	Myo3p	Myo4p	Myo5p	
	31 (27.5)	(25.3)	32 (28.6)	32 (25.9)	**Myo1p**
		37 (27.6)	50 (51.4)	37 (28.1)	**Myo2p**
			36 (26.6)	73 (75.7)	**Myo3p**
				32 (28.0)	**Myo4p**
					Myo5p

Fig. 21 (a). Phylogenetic tree for the Myo1p/Myo2p/Myo3p/Myo4p/Myo5p proteins encoded by genes of the myosin family using *S. pombe* myosin II homolog (accession no. T41522) as an outgroup. (b) Amino acid percentage identity of the myosin family.

A.

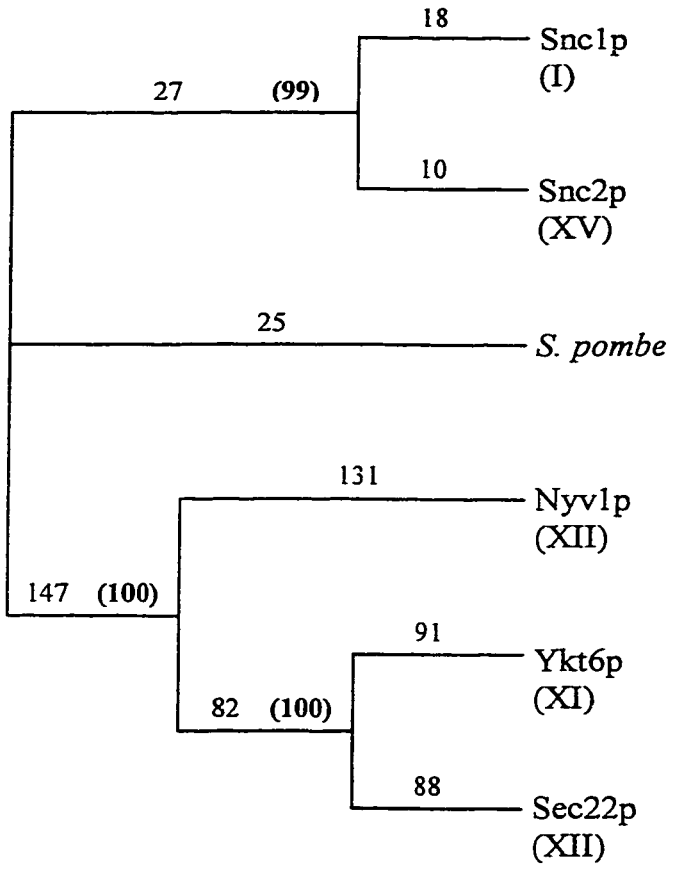


B.

Pmt1p	Pmt2p	Pmt3p	Pmt4p	Pmt5p	Pmt6p	
	35 (31.7)	33 (29.1)	32 (29.5)	54 (40.4)	33 (30.6)	**Pmt1p**
		67 (65.3)	30 (28.4)	31 (23.4)	45 (44.2)	**Pmt2p**
			(24.3)	(20.9)	44 (43.7)	**Pmt3p**
				29 (19.5)	(27.1)	**Pmt4p**
					28 (21.1)	**Pmt5p**
						Pmt6p

Fig. 22 (a). Phylogenetic tree for Pmt1/Pmt2p/Pmt3p/Pmt4p/Pmt5p/Pmt6p, palmityl-O-mannosyl transferase protein family. *S. pombe* homolog (accession no. 042933 having a probable dolichyl-phosphate-mannose protein mannosyltransferase function) is used as an outgroup.(b) Amino acid percentage identity of PMT family.

A.



B.

Snc1p	Snc2p	Nyv1p	Ykt6p	Sec22p	
	77 (77.4)	36 (28.2)	32 (9.4)	(16.2)	**Snc1p**
		35 (32.2)	29 (16.5)	(20)	**Snc2p**
			(18)	(19.6)	**Nyv1p**
				28 (23.5)	**Ykt6p**
					Sec22p

Fig. 23 (a). Phylogenetic tree for the Snc1p/Snc2p/Nyv1p/Ykt6p/Sec22p protein family using an *S. pombe* homolog (accession no. Q92356 with probable synaptobrevin function) as an outgroup. (b) Amino acid percentage identity of SNC family.

(*PMT6* and *PMT4*) lost their own homologs. *PMT1* and *PMT5* occur in tandem on chromosome IV implying they were not part of the whole-genome duplication. Two scenarios are possible. In the first, the tandem duplication occurred before the whole genome event and the homologs of *PMT1* and *PMT5* were deleted. In the second, the tandem duplication occurred after the whole-genome duplication. One ancestor of *PMT1/PMT5* was lost, while the other gave rise to the tandem pair.

In the synapobrevin family, it appears that a common ancestor gave rise to the "block 2" members *SNC1* and *SNC2* while another one gave rise to *NYV1*, *YKT6* and *SEC22*. (Fig. 23A).

3.6 Allotetraploidy vs. autotetraploidy

The whole-genome duplication event proposed by Wolfe & Shields (1997) has not yet been determined to be a result of allotetraploidy (fusion of two species) or autotetraploidy (doubling of chromosomes without a subsequent cell division). However, a method for determining whether a species is actually a hybrid, containing chromosomes from two different species has been used to reveal its evolutionary history. Tamai et al. (1998) found that some cloned *S. cerevisiae* genes hybridized strongly to *S. pastorianus* chromosomes suggesting the presence of some *S. cerevisiae*-type chromosomes in *S. pastorianus*. Weakly hybridizing *S. cerevisiae* genes to *S. pastorianus* which showed the same intensity when hybridized to *S. bayanus* chromosomes suggested that *S. pastorianus* also contained some *S. bayanus*–type chromosomes. The advantage of this technique is

that tedious collection of sequence information from multiple species is not necessary to obtain this information. However, sequence information could be used to corroborate data from such hybridization studies. For example, if the *S. cerevisiae* genome was found to be derived from an allotetraploid, each gene in a pair should show closer distances to its respective "donor" species than to each other (Fig. 24A). In fact, any species from each clade to which donor species belongs (any A's compared to any B's) would show this. Notably, the closer genetic distances between genes in *S. pastorianus* and their donor species have indeed been observed (Pederson, 1985; Petersen et al. 1987; Hansen and Kielland-Brandt, 1994; Fujii et al., 1994). Conversely, gene pairs resulting from a "self" duplication should show smaller genetic distances between the duplicated genes than to closely related species (24B).

The closest relative of *S. cerevisiae* for which a substantial sequence database is currently available is *K. lactis*. Unfortunately, this species may not provide adequate data to address this problem. It pre-dates the *Saccharomyces* genus by 50 million years (Wolfe and Shields 1997) and is though to have diverged before the *Saccharomyces* clade evolved, in which the hybridization may have occured.

Fig. 24A. Theoretical phylogenetic tree displaying the fusion of two species, "A3" and "B1" from different clades, "A" and "B", in the hypothesized allotetraploid event. "K" represents a species pre-dating the two clades, as *K. lactis* would in the case of *S. cerevisiae* . In this model, sequence data from any species in clade "A" would be closer to the "A" half of the "AB" genome, than the "A" and "B" halves of the "AB" genome would be to eachother, and the same is true for the "B" clade.

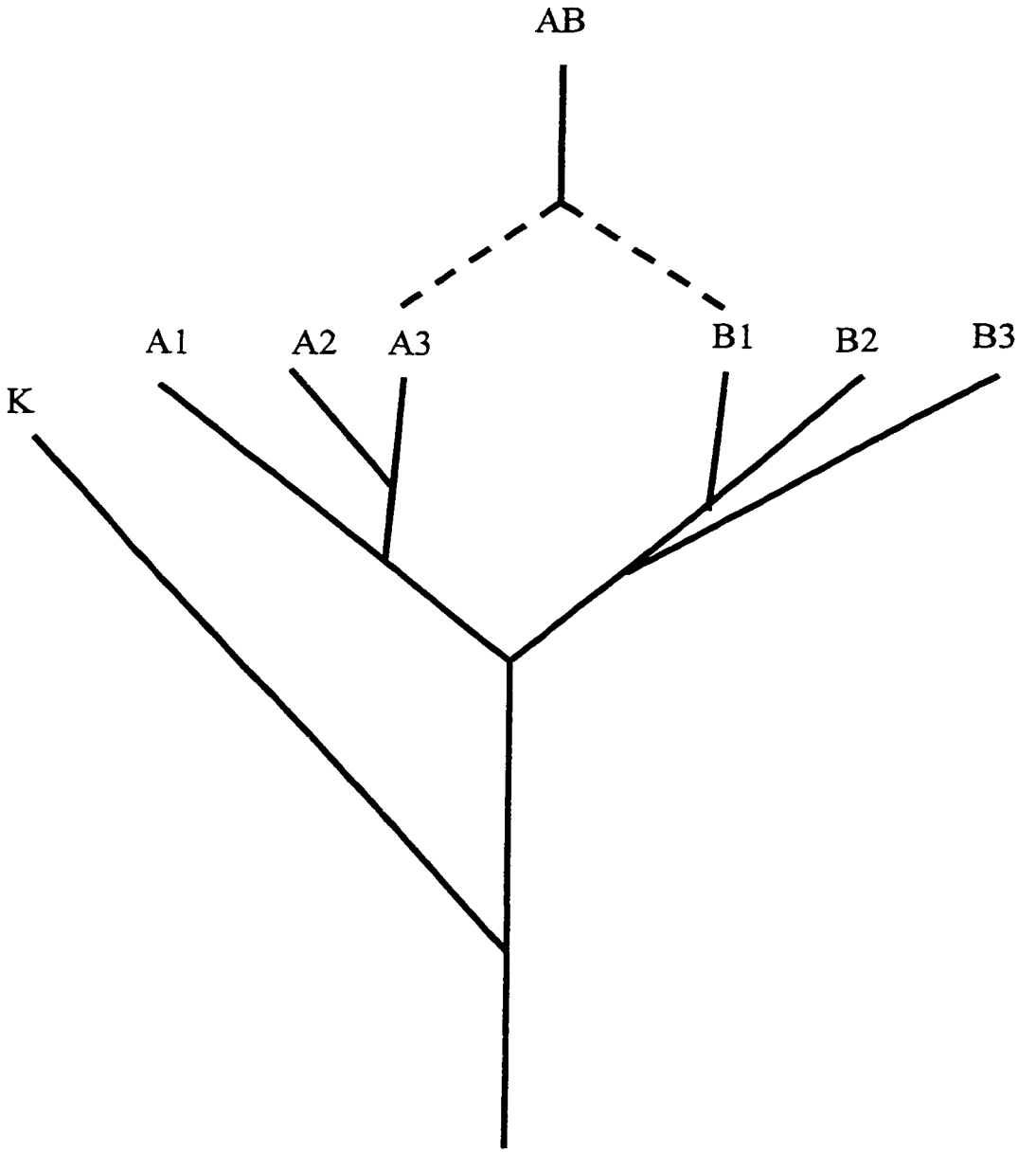
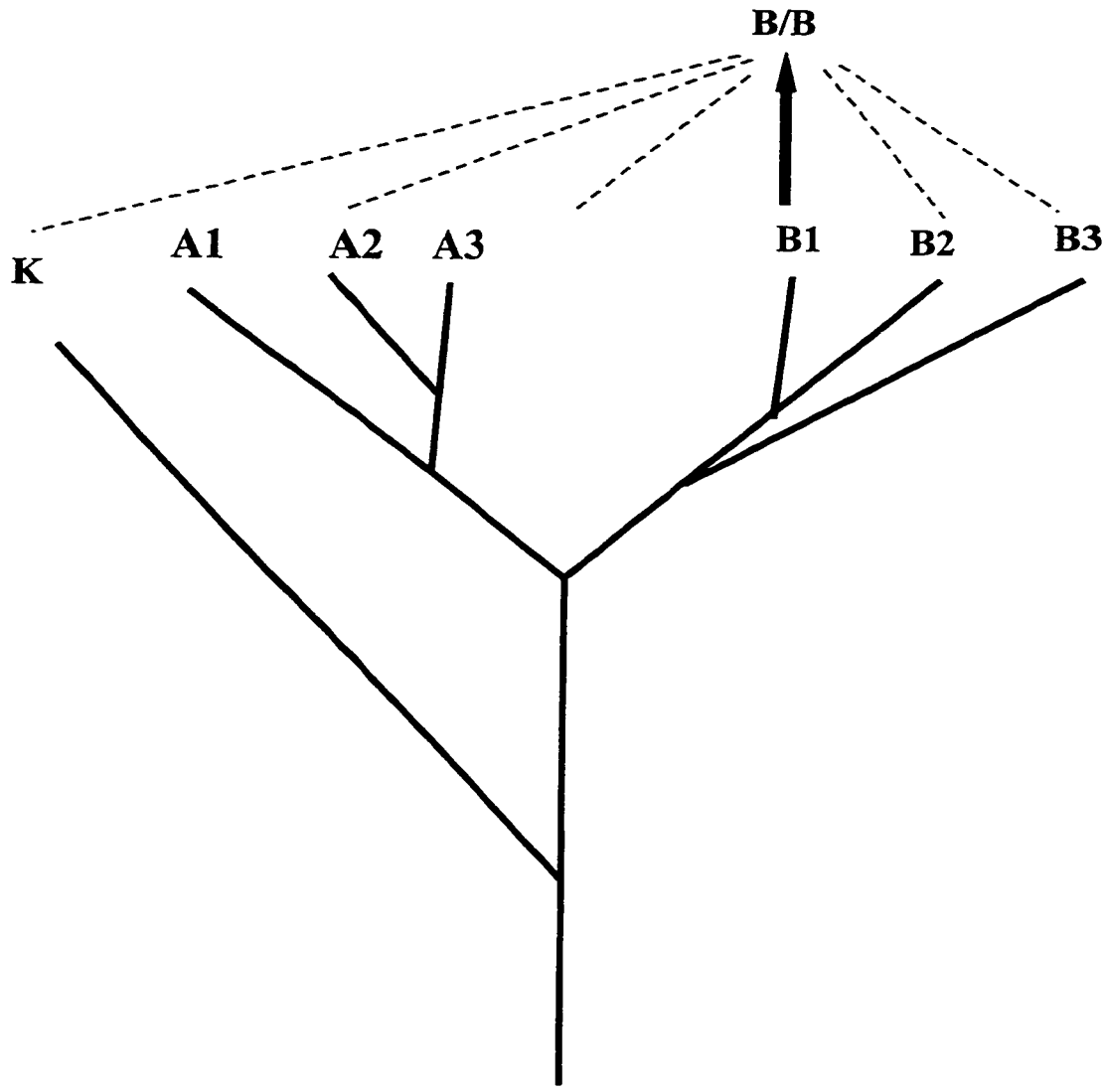


Fig. 24B. Theoretical phylogenetic tree displaying the hypothesized autotetraploid event where species "B1" undergoes a whole-genome duplication . In this model, sequence data from any species except "B1" would show a greater distance to homologs in "B/B", than would the homologs in "B/B" share themselves.



3.7 General observations

3.7.1 The significance of sequence similarity

The wide ranges in sequence identity between the gene pairs in "block 2" is notable for two reasons. First, this data suggests the percent sequence identity does not necessarily correlate with the likelihood of having functional overlap. For example, Pyk2p can not support growth on fermentable carbon sources in a strain deleted for *PYK1* despite sharing 71% amino acid identity. In contrast, *YAL056w* and *YOR371c* can individually support growth in the presence of caffeine or cycloheximide despite having only 35% identity. Secondly, the large variations in amino acid sequence identity (26% to 87% in "block 2") translate into measures of evolutionary distance that vary dramatically between gene pairs despite their apparent synchronous origin. Therefore, estimates of the time of duplication based on molecular clock models will vary with the gene pairs used.

3.7.2 Why functionally overlapping genes are maintained

It is clear how functionally divergent genes can be maintained, but a paradox arises when considering how functionally overlapping genes are maintained, as they are under some circumstances not essential. Selection could maintain both genes if they function optimally under different physiological conditions (as suspected for the glutamate biosynthesis genes), if they differed in their substrate or target specificity (ie. as do members of the *PMT* family), or if they performed similar functions in different cellular locations (*MYO2/MYO4*). In

addition, they could be used at different times in the organism's life cycle or in different gene dosages. The use of heterodimers in selecting for both copies of duplicated genes is another mechanism that is exemplified in the region examined. Heterodimeric proteins are seen in *OAF1/PIP2*, the *PMT* family, and have also been suggested to form from the two pyruvate kinase isozymes. The use of separately transcribed genes (regulatory in the first case, enzymatic in the others) forming multimers that specify targets or mediate biological pathways demonstrates an evolutionary transition towards specialization of proteins and an increase in complexity. Therefore, the use of heterodimers to increase the heterogeneity of biological processes and thus provide greater efficiency and/or specificity, is another mechanism that could be used to maintain gene duplications.

4. CONCLUSION

Functional overlap caused by the creation of homologous gene pairs during duplication is partly responsible for masking phenotypes. This can be overcome by subjecting mutants deleted for both homologous genes to phenotypic screens, although mutants containing several deletions may be required to reveal phenotypes in cases where larger gene families exist. This approach can be used to reveal phenotypes for some of the genes previously overlooked by single deletion studies.

Out of the 12 gene pairs studied, three show functional divergence: *OAF1/PIP2, PYK1/PYK2,* and *MYO4/MYO2*. Four gene pairs show functional

overlap from examining double deletions: *YAL056w/YOR371c*, *YAL031c/YOR334w*, *SNC1/SNC2*, and *PMT2/PMT3*. Two more gene pairs, *YAL053w/YOR365c* and *GDH1/GDH3*, show functional overlap between one or both paralogs and homologs outside "block 2", [*Δyal053w Δygl139w Δypl221w* and *Δgdh1 Δgdh3 Δglt1* (Avendano et al. 1997)]. No phenotypes were detected for the remaining three gene pairs (*YAL037w/YOR342c*, *YAL034c/YOR338w* and *YAL028w/YOR324c*), and they therefore remain unclassified.

Phenotypes were revealed for two groups of genes for which no members currently have phenotypic or functional data reported. Double mutant *Δyal056w Δyor371c* showed synthetic phenotypes on caffeine, cycloheximide, hydrogen peroxide and calcofluor white, which may be indicative of defects in cell surface biosynthesis. *Δyal053w Δygl139w Δypl221w* showed a synthetic lethal defect on YEPD at 30°C. Additional testing will be required to obtain further insight into the functional importance of this four member gene family.

Data from cladistic analyses of gene families showed that the gene duplications arising from the whole-genome duplication always segregated into clades as predicted by the event.

Although it has not yet been determined whether the whole-genome duplication occurred as a result of allotetraploidy or autotetraploidy, future work using interspecific chromosome hybridizations and sequence analyses looks like a promising approach to address this problem.

REFERENCES

Asubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith J.A., and Struhl, K, Eds. 1988. Current protocols in molecular biology, Vol. 2. Greene Publishing Associates, New York.

Avendano, A., Deluna, A., Olivera, H., Valenzuela, L., Gonzalez., A. 1997. GDH3 encodes a glutamate dehydrogenase isozyme, a previously unrecognized route for glutamate biosynthesis in *Saccharomyces cerevisiae.* J. Bacteriol. 179, 5594-5597.

Baudin, A., Ozier-Kalogeropoulos, O. Denouel, A. Lacroute, F. Cullin, C.1993. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae.* Nucleic Acids Res 21, 3329-3330.

Bobola, N., Jansen, R.P., Shin, T.H., and Nasmyth, K. 1996. Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. Cell 84, 699-709.

Boles, E., de Jong-Gubbels, P., and Pronk, J.T. 1998. Identification and characterization of MAE1, the *Saccharomyces cerevisiae* structural gene encoding mitochondrial malic enzyme. J. Bacteriol. 180, 2875-2872.

Boles, E., Schulte, F., Miosga, T., Freidel, K., Schluter, E., Zimmermann, F.K., Hollenberg, C.P., and Heinisch, J.J. 1997. Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in Saccharomyces cerevisiae that is catalytically insensitive to fructose-1,6-bisphosphate. J. Bacteriol. 179, 2987-2993.

Boone, C., Sommer, S.S., Hensel, A. and Bussey, H. 1990. Yeast *KRE* genes provide evidence for a pathway of cell wall β-glucan assembly. J. Cell Biol. 110, 1833-1843.

Brockerhoff, S.E., Stevens, R.C., and Davis, T.N. 1994. The unconventional myosin, Myo2p, is a calmodulin target at sites of cell growth in *Saccharomyces cerevisiae.* J. Cell. Biol. 124, 315-323.

Bui, D.M., Gregan, J., Jarosch, E. Ragnini, A. Schweyen, R. J. 1999. The bacterial magnesium transporter CorA can functionally substitute for its putative homologue mrs2p in the yeast inner mitochondrial membrane. J Biol Chem 274, 20438-20443.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., and Herskowitz, I. 1998. The transcriptional program of sporulation in budding yeast. Science 282, 699-705.

Coissac, E., Maillier, E., Netter, P. 1997. A comparative study of duplications in bacteria and eukaryotes: the importance of telomeres. Mol. Biol. Evol. 14, 1062-1074.

Cooper, T.G. 1982. Nitrogen metabolism in *Saccharomyces cerevisiae*. In *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression*, pp. 39-99. Edited by J.N. Strathern, E.W. Jones & J.R. Broach. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

DeRisi, J.L., Iyer, V.R., Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278, 680-6.

Folch, J.L., Antaramian, A., Rodriguez, L., Bravo, A., Brunner, A., Gonzalez, A. 1989. Isolation and characterization of a *Saccharomyces cerevisiae* mutant with impaired glutamate synthase activity. J. Bacteriol. 171, 6776-6781.

Fujii, T., Nagasawa, N., Iwamatsu, A., Bogaki, T., Tamai, Y. and Hamachi, T. 1994. Molecular cloning, sequence analysis, and expression of the yeast alcohol acetyltransferase gene. Appl. Environ. Microbiol. 60, 2786-2792.

Gentzsch, M., Immervoll, T., and Tanner, W. 1995. Protein O-glycosylation in *Saccharomyces cerevisiae*: the protein O-mannosyltransferases Pmt1p and Pmt2p function as a heterodimer. FEBS Lett. 377, 128-130.

Gentzsch, M. and Tanner, W. 1996. The PMT gene family: protein O-glycosylation in *Saccharomyces cerevisiae* is vital. Embo. 15, 5752-5759.

Goebl, M.G. and Petes, T.D. 1986. Most of the yeast genomic sequences are not essential for cell growth and division. Cell. 46, 983-992.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., et al. 1996. Life with 6000 genes. Science 274, 546-547.

Govindan, B., Bowser, R., and Novick, P. 1995. The role of Myo2, a yeast class V myosin, in vesicular transport. J. Cell Biol. 128, 1055-1068.

Grenson, M. and Hou, C. 1972. Ammonia inhibition of the general amino acid permease and its suppression in NADP-specific glutamate dehydrogenaseless mutants of *Saccharomyces cerevisiae*. Biochem. Biophys. Res. Comm. 48, 749-758.

Haarer, B.K., Petzold, A., Lillie, S.H., and Brown, S.S. 1994. Identification of MYO4, a second class V myosin gene in yeast. J. Cell Sci. 107, 1055-1064.

Hampsey, Michael. 1997. A review of phenotypes in *Saccharomyces cerevisiae*. Yeast 13, 1099-1133.

Hansen, J., and Kielland-Brandt M.C. 1994. *Saccharomyces carlsbergensis* contains two functional *MET2* alleles similar to homologues from *S. cerevisiae* and *S. monacensis*. Gene 140, 33-40.

Hill. J.E., Myers, A.M., Koerner., T.J., Tzagoloff, A. 1993. Yeast/E. coli shuttle vectors with multiple unique restriction sites. Yeast 2, 163-167.

Hofmann, K., Bucher P., Falquet L., Bairoch A. 1999. The PROSITE database; its status in 1999. Nucleic Acids Res. 27, 215-219.

Holzer, H., and Schneider, S. 1957. Anreicherung und Trennung einer DPN-spezifischen und einer TPN-spezifischen Glutaminosaure Dehydrogenase aus Hefe. Biochem. Z. 329, 361-367.

Immervoll, T., Gentzsch, M., and Tanner, W. 1995. PMT3 and PMT4, two new members of the Protein-O-Mannosyltransferase gene family of *Saccharomyces cerevisiae*. Yeast 11, 1345-1351.

Jansen, R.P., Dowzer, C., Michaelis, C., Galova, M., and Nasmyth, K. 1996. Mother cell-specific HO expression in budding yeast depends on the unconventional myosin Myo4p and other cytoplasmic proteins. Cell 84, 687-697.

Johnston, G.C., Prendergast, J.A., and Singer, R.A. 1991. The *Saccharomyces cerevisiae* MYO2 gene encodes an essential myosin for vectorial transport of vesicles. J. Cell. Biol. 113, 539-551.

Karpichev, I.V., Luo, Y., Marians, R.C., and Small, G.M. 1997. A complex containing two transcription factors regulates peroxisome proliferation and the coordinate induction of beta-oxidation enzymes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 17, 69-80.

Karpichev, Igor V., and Small, Gillian M. 1998. Global regulatory functions of Oaf1p and Pip2p (Oaf2p), transcription factors that regulate genes encoding peroxisomal proteins in *Saccharomyces cerevisiae*. Mol. and Cell. Biol. 18, 6560-6570.

Koll, H. Schmidt, C., Wiesenberger, G. and Schmelzer, C. 1987. Three nuclear genes suppress a yeast mitochondrial splice defect when present in high copy number. Curr. Genet. 12, 503-509.

Lalo, D., Stettler, S., Mariotte, S., Slonimski, P.P., Thuriaux, P. 1993. Une duplication fossile entre les régions centromériques de deux chromosomes chez la levure. C.R. Adac. Sci. Paris 316, 367-373.

Levin, D.E. and Bartlett-Heubusch, E. 1992. Mutants in the *S. cerevisiae PKC1* gene display a cell cycle-specific osmotic stability defect. J. Cell. Biol. 116, 1221-1229.

Lillie S.H. and S.S. Brown. 1994. Immunofluorescence localization of the unconventional myosin, Myo2p, and the putative kinesin-related protein, Smy1p, to the same regions of polarized growth in *Saccharomyces cerevisiae.* J Cell Biol 125, 825-842.

Long, R.M., Singer, R.H., Meng, X., Gonzalez, I., Nasmyt, K., and Jansen, R.P. 1997. Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. Science 277, 383-387.

Luo, Yi, Karpichev, Igor V., Khanski, Ronald A., and Small, Gillian M. 1996. Purification, identification, and properties of a *Saccharomyces cerevisiae* oleate-activated upstream activating sequence-binding protein that is involved in the activation of *POX1*. J. of Biol. Chem. 271, 12068-12075.

Lussier, M., Gentzsch, M., Sdicu, A.M., Bussey., H., and Tanner W. 1995. Protein O-glycosylation in yeast; the PMT2 gene specifies a second protein O-mannosyltransferase that functions in addition to the PMT1-encoded activity. J. Biol. Chem. 270, 2770-2775.

Mackiewicz, P., Kowalczuk, M., Gierlik, A., Dudek, M.R., and Cebrat, S. 1999. Origin and properties of non-coding ORFs in the yeast genome. Nucleic Acids Res. 27, 3503-3509.

McCusker, J.H. and Haber, J.E. 1988. Cycloheximide-resistant temperature-sensitive lethal mutations of *Saccharomyces cerevisiae*. Genetics 119, 303-315.

Melnick, L., Sherman, F. 1993. The gene clusters ARC and COR on chromosomes 5 and 10, respectively, of *Saccharomyces cerevisiae* share a common ancestry. J. Mol. Biol. 233, 372-388.

Mewes, H.W., Albermann, K., Bahr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K., Kleine, K., Maierl, A., Oliver, S.G., Pfeiffer, F., Zollner, A. 1997. Overview of the yeast genome. Nature 387, Suppl. 7-65.

Mewes, H.W., Heumann, K., Kaps, A., Mayer, K., Pfeiffer, F., Stocker, S., Frishman, D. 1999. MIPS: a database for protein sequences and complete genomes. Nucleic Acids Res. 27:44-48.

Mushegian, Arcady R., Garey, James R., Martin, Jason, Liu, Leo X. 1998. Large-scale taxonomic profiling of eukaryotic model organisms: A comparison of orthologous proteins encoded by the human, fly, nematode and yeast genomes. Genome Res. 8:590-598.

Nakai, K., and Kanehisa, M. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14, 897-911.

Ohno, S. 1970. *Evolution by gene duplication.* Springer-Verlag, Berlin, New York.

Oliver, S.G. 1996. From DNA sequence to biological function. Nature 379, 597-600.

Paravicini, G., Cooper, M., Friedli, L., Smith, D.J., Carpentier, J.-L., Klis, L.S. and Payton, M.A. 1992. The osmotic integrity of the yeast cell requires a functional *PKC1* gene product. Mol. Cell. Biol. 12, 4896-4905.

Pearson, W.R. and Lipman, D.J. 1988. Improved tools for iological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444-2448.

Pedersen, M.B. 1985. DNA sequence polymorphisms in the genus *Saccharomyces*, II. Analysis of the genes *RDN1*, *HIS4*, *LEU2* and *Ty* transposable elements in Carlsberg, Tuborg and 22 Bavarian brewing strains. Carlsberg Res. Commun. 50, 263-272.

Petersen, J.G.L., Nilsson-Tillgren, T., Kielland-Brandt, M.C., Gjermansen, C. and Holmberg, S. 1987. Structural heterozygosis at genes *ILV2* and *ILV5* in *Saccharomyces carlsbergensis*. Curr. Genet. 12, 167-174.

Philippsen, P., Kleine, K., Pohlmann, R., Dusterhoft, A., Hamberg, K., et al. 1997. The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIV and its evolutionary implications. Nature 387, 93-98.

Protopopov, V., Govindan, B., Novick, P., and Gerst, J.E. 1993. Homologs of the synaptobrevin/VAMP family of synaptic vesicle proteins function on the late secretory pathway in *S. cerevisiae*. Cell 74, 855-861.

Ram, A.F., Wolters, A., Ten Hoopen, R., Klis, F.M. 1994. A new approach for isolating cell wall mutants in *Saccharomyces cerevisiae* by screening for hypersensitivity to calcofluor white. Yeast 10, 1019-1030.

Rottensteiner, H., Kal, A.J., Filipits, M., Binder, M., Hamilton, B., Tabak, H.F., and Ruis, H. 1996. Pip2p: a transcriptional regulator of peroxisome proliferation in the yeast *Saccharomyces cerevisiae*. EMBO J. 15, 2924-2934.

Rottensteiner, Hanspeter, Kal, Arnoud J., Hamilton, Barbara, Ruis, Helmut, and Tabak, Henk F. 1997. A heterodimer of the Zn_2Cys_6 transcription factors Pip2p and Oaf1p controls induction of genes encoding peroxisomal proteins in Saccharomyces cerevisiae. Eur. J. Biochem. 247, 776-783.

Seioghe, C., and Wolfe, K.H. 1999. Updated map of duplicated regions in the yeast genome. Gene 238, 253-261.

Shaw, D.R., Richter, H., Giorda, R., Ohmachi, T., and Ennis, H.L. 1989. Nucleotide sequences of *Dictyostelium discoideum* developmentally regulated cDNAs rich in (AAC) imply proteins that contain clusters of asparagine, glutamine or threonine. Mol. Gen. Genet. 218, 453-459.

Smith, M.M. 1987. Molecular evolution of the *Saccharomyces cerevisiae* histone gene loci. J. Mol. Evol. 24, 232-259.

Storms, R.K., Wang, Y., Fortin, N., Hall, J., Vo, D.H., Zhong, W., Downing T., Barton, Arnold B., Kaback, D.B., Su, Y., and Bussey, H. 1997. Analysis of a 103 kbp cluster homology region from the left end of *Saccharomyces cerevisiae* chromosome I. Genome 40, 151-164.

Tamai, Y., Momma, T., Yoshimoto, H., and Kaneko, Y. 1998. Co-existence of two types of chromosome in the bottom fermenting yeast, *Saccharomyces pastorianus*. Yeast, 14, 923-933.

Tempest, D.W., Meers, J.L., and Brown, C.M. 1970. Synthesis of glutamate in *Aerobacter aerogenes* by a hitherto unknown route. Biochem. J. 117, 405-407.

Thatcher, J.W., Shaw, Janet M., and Dickinson, W.J. 1998. Marginal fitness contributions of nonessential genes in yeast. Proc. Natl. Acad. Sci. USA. 95, 253-257.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. 1997. The Clustal_X windows interface; flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882.

Ushinsky, S.C., Bussey, H., Ahmed, A.A., Wang, Y., Friesen, J., Storms, R.K. 1997. Functional analysis of a 38 kilobase region on chromosome XVI in Saccharomyces cerevisiae. Genes Funct. 4, 273-84.

Wach, A., Brachat, A., Pohlmann, R., Philippsen, P. 1994. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. Yeast 10, 1793-1808.

Waldherr, M., Ragnini, A., Jank, B., Teply, R., Wiesenberger, G., and Schweyen, R.J. 1993. A multitude of suppressors of group II intron-splicing defects in yeast. Curr. Genet. 24, 301-306.

Wiesenberger, G., Waldherr, M., and Schweyen, R.J. 1992. The nuclear gene MRS2 is essential for the excision of group II introns from yeast mitochondrial transcripts *in vivo*. J. Biol. Chem. 267, 6963-6969.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K. et al. 1999. Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. Science 285, 901-906.

Wolfe, K.H., and Lohan, A.J. 1994. Sequence around the centromere of *Saccharomyces cerevisiae* chromosome II: similarity of CEN2 to CEN4. Yeast 10, Suppl. A:S41-46.

Wolfe, Kenneth H. and Shields, Denis C. 1997. Molecular evidence for an ancient duplication of the yeast genome. Nature 387, 708-713.

APPENDIX

```
CLUSTAL X (1.64b) multiple sequence alignment

yal031          MVDVQKRKKLLAKAAASASIPAIKGSVPLDSYDIKIIQYKNALYKLNELNRLLNVLVPHL 60
yor334          ----MNRR-LLVR-SISCFQPLSR--ITFGRPNTPFLR-K---YADTSTAANTNSTI--- 45
                     :*: **.: : *.  *  :  :.:.  :  ::: *   *  ..     *  :

yal031          KKKRDNDESYKIIPLVNFILSLCEGPIFNVSPVLAKRYHLLCRFQLIKLSEVQQRLSTNF 120
yor334          --LRKQLLSLKPISASDSLFISCT--VFNS-----------KGNIISMSEKFPKWS--F 87
                   *.:  * * *.  : ::  *   :**             : ::*.:**   : *  *

yal031          IDVEGWMFPEEVPLDHYKSCIYNNSLQWKILNSLSCIAQNAIKIYNAKLRQILLERDAYK 180
yor334          L-TEHSLFPR----DLRK--IDNSSI--DIIPTIMC-KPNCIVI-NLLHIKALIERDKVY 136
                : .*  :**.    *  *  * *.*:  .*: :: *   *.* * *    : *:***

yal031          ARSLPFDTSIIEDLLNPVEMTLILDLAVLINDPVRDKSTHSFYKLQWQVMEKLNSCVHSK 240
yor334          V----FDT------TNP---SAAAKLSVLMYD---LESKLSSTKNNSQFYE------HRA 174
                .    ***       **   :   .*:**: *    :*. *  * : *. *      *

yal031          IFPILRTYYNQLQKFSETRPTSLSNLQKDLPHWEWTLHRIYTFHLRVFSVLCVIISFSRQ 300
yor334          LESIFINVMSALETDFKLHSQICIQILNDLEN---EVNRLKLRHL-------LIKSKDLT 224
                : .*: .  . *:.  : :.    :: :** :    ::*:   **       :* * .

yal031          IFLPNKQHFLDIKTRLSSENVYHYDLIICELMALLSPECDDVTALFELQENLKFWTQTAR 360
yor334          LFYQKTLLIRDLLDELL-ENDD--DLANMYLTVKKSPK-DNFSDLEMLIETY--YTQ--- 275
                :*  :.  : *:  .*  **    **    * .  **: *:.: *  * *.   :**

yal031          TDNNSSRTPIFHLQPGLVVELFNNHICKIIPKLRSIMGLLSNWMDCWKYIEKNYKTFDET 420
yor334          CDEYVQQSESL-IQDIKSTEEIVNIILDANRNSLMLLELKVTIYTLGFTVASVLPAFYGM 334
                 *:  .::  : :*    .* : * * .   :   :: *  .       : .   :*

yal031          NDLRENLKEKLERDKALYLEVKNAKSKLKKKPSITKLPASSSPSPSPTSSASPSRQASLE 480
yor334          N--LKNFIEESEWGFTSVAVFSIVSALYITKKNFNSLRSVTKMTMYPNS---P------- 382
                *   :*: *: * . :    .. ..:   .* .:..* : :. :  *.*   *

yal031          SIRTRARAHLASNSSRSPSVSPVRTTFNNKNAETKKSVVSPEKRKLINGRRPRSSSLQSY 540
yor334          -----------ANSSVYPKTS-ASIALTNKLKRRRKWWKSTKQR---LGVLLYGS---SY 424
                          :***  *..* .  ::.**  . :*   *.::*    *    .*   **

yal031          TNKQQTSYLNSTRHPSIAPPSKLNNQRSNSLQSSTMTLNQKIVQDTVRHLMNKSASTPNP 600
yor334          TNKANLS----------------NN----------------------------------- 433
                *** : *                **

yal031          SASSSLAPSPKVSSINNTSSGKSSSTLIANSSDTLAIETLTLDPESNSSELSIKRVRFAG 660
yor334          -------------KIN---KG-----------------------------------FSK 441
                             .**   .*                                     *:

yal031          VPPMT-EAENPKPTKVGWYKKPAVLHYPPIPASAMIKPLQHKSKYNTLRQEEGFTFRKSL 719
yor334          VKKFN--MEN-------------------------DIK--------N----------KQN 456
                *  :.   **                          **        *          *.

yal031          RDGLEWENGESGSETTMMPFGIEIKESTGHRIASKIRSKLR------------------- 760
yor334          RD-MIWK-----------WLIEDK---------KN------------------------ 470
                ** : *:            : ** *         *

yal031          ------------------------------------------------------------
yor334          ------------------------------------------------------------


yal031          ------------------------------------------------------------
yor334          ------------------------------------------------------------


yal031          -----------------
yor334          -----------------
```

```
CLUSTAL X (1.8) multiple sequence alignment


yal037          MDMEIEDSSPIDDLKLQKLDTNVYFGPCEILTQPILLQYENIKFIIGVNLSTEKIASFYT
yor342          MTILEELN---DSSIPQRLDNHIFFGSVHSLTHTDFLVENNIRFFINVDLSTELISHIYH
                * :  * .   *.   *:**.:::**. . **:. :*  :**:*:*.*:**** *: :*

yal037          QYFRNSNSVVVNLCSPTTAAVATKKAAIDLYIRNNTILLQKFVG--QYLQMGKKIKTSLT
yor342          EVRSKFAHEIVIVNIDNDSQIPIESDLVRSFHWHNTSLLQQLIHHLDFLSGINNHGEPLT
                :   :    :* :   . : :. :.  :  :  :** ***:::   ::*.  ::   .**

yal037          --------------QAQTDTIQSLP--------QFCNSNVLSGEPLVQYQAFNDLLALFK
yor342          PPPESHYRNAYVQFDHPSDSVSILDKLLYGNKSEYSRTNIFQVTNEAKFQVFNDLITIFK
                              :  :*::. *         :: ..:*::.    .::*.****:::**

yal037          SFSHFG-----NILVIS-SHSYDCALLKFLISRVMTYYPLVTIQDSLQYMKATLNISIST
yor342          YSIAQGGNTNSNILVLSENGSTDENLISLLMSTVLKENPTFNVYQALQFVKSIAVIPDTV
                     *     ****:* . * *  *:.:*:* *: .  * ..: ::**::*:   *. :.

yal037          SDE----------FDILNDKE--LWEFG-QTQEILK--RRQTSSVKRRCVNLPENSTIDN
yor342          RDEKILWVTGFINYQELIKKNEMYWGLGSQKGRKLTSFASPISKVERKQRRRDDQNIMRS
                 **          :: * .*:   * :* *. . *.      *.*:*:  .  ::. : .

yal037          RMLMG---------TTKRGRF-
yor342          KLPQQRQNPFCSTERPKRARCD
                ::              .**.*
```