Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*

Nancy Burns,[1] Brian Grimwade,[1] Petra B. Ross-Macdonald,[1] Eui-Yul Choi, Karin Finberg, G. Shirleen Roeder, and Michael Snyder[2]

Department of Biology, Yale University, New Haven, Connecticut 06520-8103 USA

We have developed a large-scale screen to identify genes expressed at different times during the life cycle of *Saccharomyces cerevisiae* and to determine the subcellular locations of many of the encoded gene products. Diploid yeast strains containing random *lacZ* insertions throughout the genome have been constructed by transformation with a mutagenized genomic library. Twenty-eight hundred transformants containing fusion genes expressed during vegetative growth and 55 transformants containing meiotically induced fusion genes have been identified. Based on the frequency of transformed strains producing β-galactosidase, we estimate that 80–86% of the yeast genome (excluding the rDNA) contains open reading frames expressed in vegetative cells and that there are 93–135 meiotically induced genes. Indirect immunofluorescence analysis of 2373 strains carrying fusion genes expressed in vegetative cells has identified 245 fusion proteins that localize to discrete locations in the cell, including the nucleus, mitochondria, endoplasmic reticulum, cytoplasmic dots, spindle pole body, and microtubules. The DNA sequence adjacent to the *lacZ* gene has been determined for 91 vegetative fusion genes whose products have been localized and for 43 meiotically induced fusions. Although most fusions represent genes unidentified previously, many correspond to known genes, including some whose expression has not been studied previously and whose products have not been localized. For example, Sec21–β-gal fusion proteins yield a Golgi-like staining pattern, Ty1–β-gal fusion proteins localize to cytoplasmic dots, and the meiosis-specific Mek1/Mre4–β-gal and Spo11–β-gal fusion proteins reside in the nucleus. The phenotypes in haploid cells have been analyzed for 59 strains containing chromosomal fusion genes expressed during vegetative growth; 9 strains fail to form colonies indicating that the disrupted genes are essential. Fifteen additional strains display slow growth or are impaired for growth on specific media or in the presence of inhibitors. Of 39 meiotically induced fusion genes examined, 14 disruptions confer defects in spore formation or spore viability in homozygous diploids. Our results will allow researchers who identify a yeast gene to determine immediately whether that gene is expressed at a specific time during the life cycle and whether its gene product localizes to a specific subcellular location.

[*Key Words*: Gene expression; protein localization; *S. cerevisiae*; fusion protein]

Received January 5, 1994; accepted in revised form March 14, 1994.

Eukaryotic genomes are estimated to contain 6,000–100,000 genes. Even in the best characterized organisms, the functions of the vast majority of genes are unknown. Relatively little information is available concerning the fraction of the genome that is expressed in particular cell types and the cellular processes in which the gene products participate. In an attempt to gain further information concerning the structure and function of eukaryotic genomes, efforts have begun to sequence the genomes of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and humans.

S. cerevisiae is an ideal model organism for the study of eukaryotic genomes. Basic cellular processes in yeast are similar in most respects to those of other eukaryotic organisms, yet the haploid nuclear genome consists of only 14 million bp that encode 6000–8000 genes (Lauer et al. 1977; Link and Olson 1991; Olson 1991). In general, yeast genes do not contain introns, and those introns that do exist are small (average size, 300–500 bp) and usually located near the 5′ end of the gene (Woolford 1989). Hence, open reading frames (ORFs) can be identified easily. Finally, the functions of individual ORFs can be evaluated readily, as it is easy to create mutations in cloned genes and to substitute the altered versions for the genomic copies (Rothstein 1991).

Excluding the rDNA, ~25% of the yeast genome has been sequenced thus far (estimated from sequences in GenBank), and efforts to sequence the entire genome systematically have already been initiated (Oliver et al.

[1]These authors contributed equally to this work.
[2]Corresponding author.

1992; B. Dujon and A. Goffeau, pers. comm.). From the existing sequence data, it is clear that the yeast genome is densely packed with ORFs. For example, chromosome III, which is 315,000 bp in length, contains 182 ORFs that are >100 bp in length (Oliver et al. 1992). Of these, 37 ORFs correspond to known genes, and the amino acid sequences predicted by another 29 ORFs are similar to sequences in current data bases. For most predicted proteins, no significant sequence similarity is observed. Whether or not similarity is detected, the information obtained directly from sequence analysis is relatively limited. In most cases, it remains unknown whether the gene product is important at a particular stage of the life cycle, where in the organism the protein acts, or in what cellular process the gene product participates.

A variety of different approaches have been used to characterize genes and their products. In several organisms, fusions to the bacterial *lacZ* gene have been extremely useful in identifying genes expressed at different times during the life cycle (e.g., *Drosophila*; O'Kane and Gehring 1987) or under different growth conditions (e.g., *Escherichia coli*; Casadaban and Cohen 1979; Casadaban et al. 1983). Fusions to *lacZ* have also been used widely for subcellular localization of gene products by indirect immunofluorescence with antibodies to β-galactosidase (β-gal, see below). Another approach for analyzing gene function is to disrupt the gene *in vivo* and determine the resulting phenotype (Rothstein 1991).

Described below is an insertional mutagenesis scheme for the large-scale characterization of genes in *S. cerevisiae*. This scheme permitted the rapid analysis of gene expression during vegetative growth and meiosis as well as the subcellular localization of the products of many yeast genes. The scheme also allowed the analysis of the phenotypes resulting from disruption of the mutagenized genes. Data on gene expression, protein localization, and disruption phenotypes are presented for a large number of known and novel genes.

Results

Generation of yeast strains with lacZ *inserted at random locations*

Yeast strains carrying *lacZ* fusion genes were constructed according to the scheme outlined in Figure 1. First, a yeast genomic library was constructed in a vector suitable for transposon mutagenesis in bacteria. The library was mutagenized in *E. coli* with a mini-Tn*3*::*LEU2* transposon containing *lacZ*-coding sequences (Seifert et al. 1986a) to generate a large number of independent *lacZ* insertions. The transposon contains 38-bp terminal repeats; the ORF in one repeat extends into the adjacent *lacZ*-coding region. The *lacZ* gene lacks both a promoter and an initiator ATG codon; thus, β-gal production in yeast depends on insertion in-frame into the coding region of an expressed gene. The mutagenized yeast DNA sequences were released from vector DNA by digestion with a restriction enzyme and introduced into a diploid yeast strain by transformation and selection for the

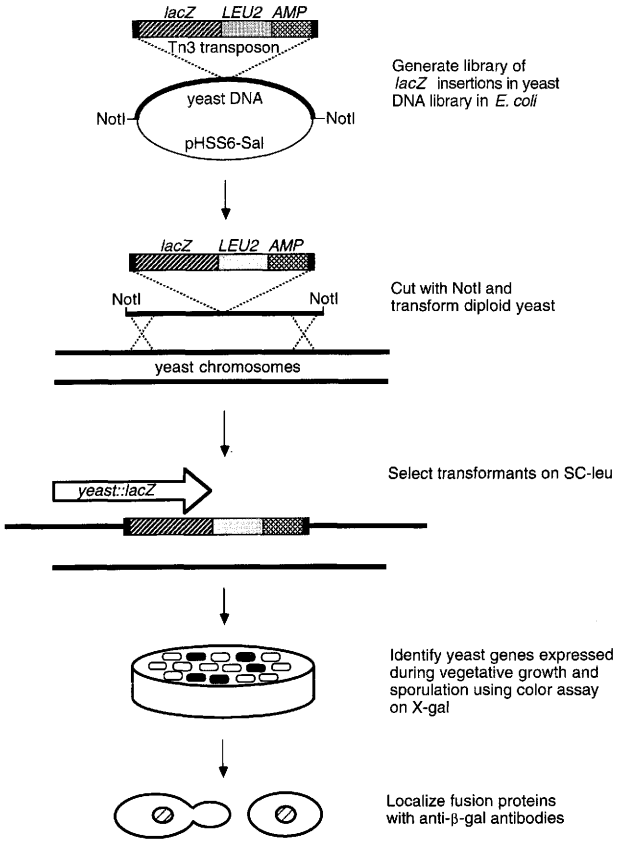

Figure 1. Overall scheme for generation and analysis of yeast::*lacZ* fusion genes.

LEU2 marker in the transposon. Transformation is expected to result from recombination between yeast DNA sequences flanking the transposon and homologous genomic sequences, resulting in replacement of the genomic copy by the mutagenized version (Fig. 1). For most of our studies, a yeast strain lacking the 2μ circle plasmid was used; thus, most transformants are expected to contain insertions into chromosomal DNA. DNA gel blot hybridization analysis of 68 random transformants revealed that 60 contain one copy of the transposon and 8 (11.8%) contain two copies. Thus, most yeast strains carry a single copy of the transposon inserted into genomic DNA.

Identification of yeast genes expressed during vegetative growth

To identify genes expressed during vegetative growth, independent transformants were patched onto synthetic complete medium lacking leucine (SC−leu) and then assayed for β-gal activity using plates containing X-gal (Xie et al. 1993). Of 20,250 patches analyzed, 2800 reproducibly turn blue. Assuming that the *lacZ* transposon inserted randomly throughout the genome and that 11.8% of the transformants contain two copies of the transposon, these data indicate that 74.2% of the ge-

nome consists of ORFs that are expressed during vegetative growth. Because 1–2 Mbp of rDNA is present in haploid yeast DNA (Link and Olson 1991) and the yeast genome is 13.5–14.5 Mbp in size (Olson 1991), 80–86% of the nonribosomal genomic DNA encodes proteins produced during vegetative growth. This figure is comparable to the fraction of chromosome III DNA that is coding (Oliver et al. 1992), suggesting that most yeast ORFs are expressed, at least at some level, during vegetative growth. The high fraction of transformants producing β-gal further suggests that the sensitivity of the patch tests is sufficient to detect almost all of the transformants carrying fusion genes expressed during vegetative growth.

Subcellular localization of fusion proteins

The subcellular locations of 2373 fusion proteins were determined by indirect immunofluorescence using rabbit anti-β-gal antibodies. To analyze a large number of yeast strains, procedures were developed for growing and processing cells in the wells of microtiter dishes, as described in Materials and methods. As a positive control for the immunofluorescence experiments, cells were stained simultaneously with a rat monoclonal antibody that recognizes tubulin (Kilmartin et al. 1982) and with the DNA-binding dye Hoechst 33258, which reveals the location of the nuclear and mitochondrial DNAs (Snyder and Davis 1988). Cells at different stages of the cell cycle (i.e., unbudded cells, and cells with small and large buds) were examined for each transformant.

A summary of the immunofluorescence results is presented in Table 1, and examples of different staining patterns are shown in Figure 2. For 68% of the transformants examined, staining above background was detected. The majority, 58% of the total, exhibit general cytoplasmic staining (e.g., Fig. 2i) that is uniform, granular, fibrous, or finely speckled. Ten percent of the fusion proteins reproducibly localize to a discrete subcellular location. One of the major classes (3% of the total) consists of fusion proteins that localize throughout the nucleus (e.g., Fig. 2a); these usually display a uniform fibrous or granular staining pattern. Other patterns include localization to the spindle pole body, microtubules, mitochondria (e.g., Fig. 2c), or the nucleolus (e.g., Fig. 2b, as evidenced by double staining with anti-nucleolar antibodies; Yang et al. 1989; H. Friedman, C. Copeland, and M. Snyder, unpubl.). Two fusion proteins were identified that localize to bud tips, reminiscent of Spa2–β-gal fusions (C. Costigan and M. Snyder, unpubl.). Two other fusion proteins localize to the nuclear rim. For one of these, some cytoplasmic staining is also detected, consistent with localization to the endoplasmic reticulum (e.g., Fig. 2h).

In addition to mitochondrial staining, a variety of other punctate patterns of cytoplasmic staining were observed (e.g., Fig. 2d–f). Some fusion proteins localize to a single dot per cell, others to 3–10 dots per cell (similar to Sec7 and Kex2 Golgi-staining patterns; Franzusoff et al. 1991; Redding et al. 1991), and others are present in 50–100 dots per cell. One example in which the dots are usually present at the cell periphery is shown in Figure 2f.

Table 1. *Immunofluorescence patterns observed*

Pattern	Number
Nuclear	
general	72
nucleolus	4
large dot on edge of nucleus	4
Nuclear rim/endoplasmic reticulum	2
Mitochondrial	
all	21
clustered	3
Cytoplasmic spots (punctate)	
1–10 dots	58
10–100 dots	64
Cell periphery	9
Bud tip/site of cytokinesis	2
Spindle pole body	2
Microtubules	1
Thick cytoplasmic fibers	3
General cytoplasmic	
uniform	697
finely speckled	327
granular	195
fibrous	158
Background	751
Total	2373

Fusions in all of the categories (except General cytoplasmic and Background) have been tested more than once. Transformants displaying cytoplasmic spots represent a very heterogeneous category; for presentation purposes, they are divided into two classes, those containing 1–10 dots and those containing 10–100 dots.

A number of the localization patterns observed have not been described previously. For example, three fusion proteins localize to thick cytoplasmic fibers (not shown), and three other fusion proteins localize to mitochondria that are clustered, but not to isolated mitochondria (Fig. 2g). It is likely that these latter fusions induce the clustering of mitochondria, as clusters are not evident in the untransformed strain (see below). In summary, many fusion proteins localize to discrete sites in the cell and a wide variety of different staining patterns are observed.

Sequence analysis of fusion genes

To determine the identity of some of the fusion genes, the DNA immediately adjacent to *lacZ* was cloned and sequenced using the strategy outlined in Figure 3. Briefly, a plasmid marked with the *URA3* gene was integrated into the transposon by recombination between the plasmid- and transposon-borne copies of the ampicillin-resistance gene. Yeast DNA was then cut with a restriction enzyme that releases the bacterial origin of DNA replication, *lacZ*, and the adjacent yeast DNA sequences as a linear fragment; this fragment was then circularized and recovered in bacteria. The sequences

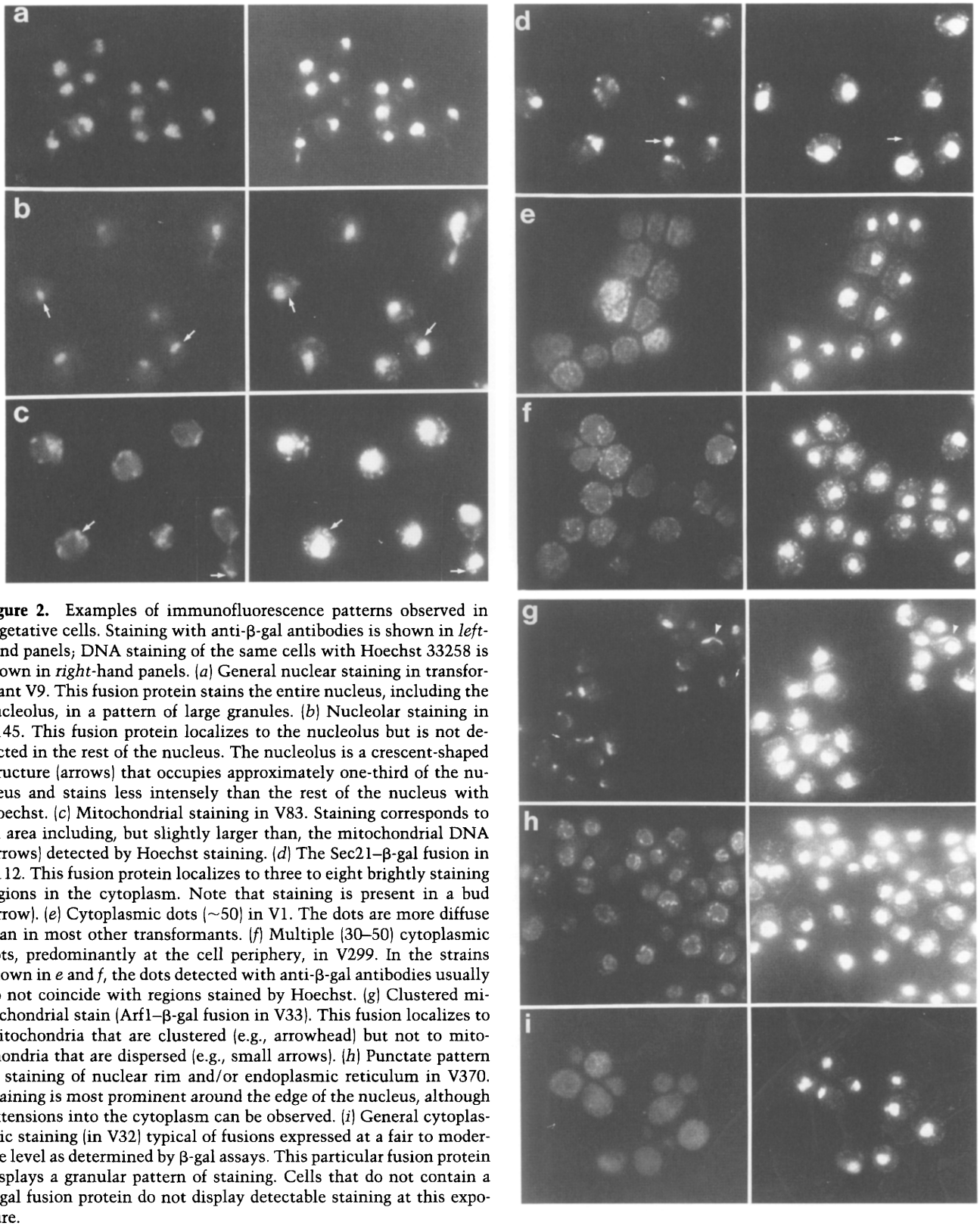

Figure 2. Examples of immunofluorescence patterns observed in vegetative cells. Staining with anti-β-gal antibodies is shown in *left*-hand panels; DNA staining of the same cells with Hoechst 33258 is shown in *right*-hand panels. (*a*) General nuclear staining in transformant V9. This fusion protein stains the entire nucleus, including the nucleolus, in a pattern of large granules. (*b*) Nucleolar staining in V145. This fusion protein localizes to the nucleolus but is not detected in the rest of the nucleus. The nucleolus is a crescent-shaped structure (arrows) that occupies approximately one-third of the nucleus and stains less intensely than the rest of the nucleus with Hoechst. (*c*) Mitochondrial staining in V83. Staining corresponds to an area including, but slightly larger than, the mitochondrial DNA (arrows) detected by Hoechst staining. (*d*) The Sec21–β-gal fusion in V112. This fusion protein localizes to three to eight brightly staining regions in the cytoplasm. Note that staining is present in a bud (arrow). (*e*) Cytoplasmic dots (~50) in V1. The dots are more diffuse than in most other transformants. (*f*) Multiple (30–50) cytoplasmic dots, predominantly at the cell periphery, in V299. In the strains shown in *e* and *f*, the dots detected with anti-β-gal antibodies usually do not coincide with regions stained by Hoechst. (*g*) Clustered mitochondrial stain (Arf1–β-gal fusion in V33). This fusion localizes to mitochondria that are clustered (e.g., arrowhead) but not to mitochondria that are dispersed (e.g., small arrows). (*h*) Punctate pattern of staining of nuclear rim and/or endoplasmic reticulum in V370. Staining is most prominent around the edge of the nucleus, although extensions into the cytoplasm can be observed. (*i*) General cytoplasmic staining (in V32) typical of fusions expressed at a fair to moderate level as determined by β-gal assays. This particular fusion protein displays a granular pattern of staining. Cells that do not contain a β-gal fusion protein do not display detectable staining at this exposure.

fused to *lacZ* were determined using a primer complementary to sequences in the transposon.

Sequence analysis was performed for 90 strains that contain a single chromosomal copy of the insertion and

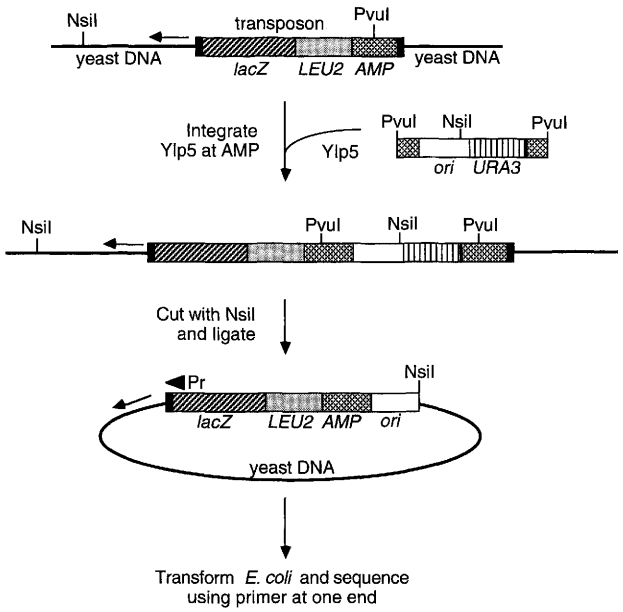

Figure 3. Scheme for plasmid rescue and sequencing.

whose β-gal fusion proteins localize to discrete sites within the cell; sequencing was also carried out for one strain that contains a fusion gene on the 2μ circle plasmid. Thirty-one of the fusions correspond to previously known genes or ORFs (Table 2), and another five encode proteins that are homologous, but not identical, to amino acid sequences present in the data bases ($P < 10^{-2}$). For the vast majority of fusion genes, the sequence is novel. A summary of these results is presented in Table 2.

Of the 22 strains that contain nuclear fusion proteins, two insertions are in genes known to encode nuclear proteins: histone H2A (Hall et al. 1984) and the 2μ circle Rep1 protein (Wu et al. 1987). Two other nuclear fusion proteins correspond to gene products involved in nuclear transcription, but whose localization has not been reported (Ada2, Berger et al. 1992; Bdf1, A. Sentenac and B. Seraphin, pers. comm.). Four mitochondrial fusion proteins correspond to genes known to be involved in mitochondrial function, including *ATP1* (Takeda et al. 1986), *HEM15* (Labbe-Bois 1990), and *NUC1* (Liu and Dieckmann 1989); two others are in known genes, *RNA12* (Liang et al. 1992) and *ARO3* (Paravicini et al. 1988), whose products have not been localized previously. The remainder of the nuclear and mitochondrial proteins are encoded by previously unidentified genes.

Sequences were also determined for strains whose fusion proteins display a punctate pattern of cytoplasmic staining. A variety of known genes were identified (see Table 2), but none of these correspond to genes for which the precise subcellular location of the encoded protein has been determined previously by immunofluorescence. However, one gene corresponds to *SEC21* whose product is involved in protein transport to or within the Golgi apparatus (Hosobuchi et al. 1992). The Sec21–β-gal fusion protein localizes to approximately four to eight cytoplasmic dots per cell (Fig. 2d), consistent with localization to the Golgi (Frazusoff et al. 1991; Redding et al. 1991). Two fusion proteins that localize to ~50 dots/cell result from insertions in Ty elements. One fusion is located in the TyA reading frame, which encodes the putative Gag protein, a component of virus-like particles (Boeke and Sandmeyer 1991). The other fusion lies in the TyB ORF, in a region presumed to encode a protease that is also present in Ty particles. The location of Ty particles in wild-type yeast strains is not known, but cytoplasmic, virus-like particles have been detected in strains overexpressing a Ty element (Garfinkel et al. 1985). The number of spots detected with anti-β-gal antibodies is consistent with the number of virus-like particles expected per cell (J. Boeke, pers. comm.).

One fusion protein that localizes to one to four spots at the cell periphery corresponds to Cap/Srv2. The Cap/Srv2 gene product stimulates adenylate cyclase (Fedor-Chaiken et al. 1990; Field et al. 1990). Adenylate cyclase and its activators, Cdc25 and Ras, appear to reside at the plasma membrane (Willumsen et al. 1986; Deschenes and Broach 1987; Mitts et al. 1990).

One of the fusion proteins that localizes to clustered mitochondria and may be responsible for the clustering corresponds to the *ARF1* gene product (Sewell and Kahn 1988). The Arf1 protein is thought to be involved in secretion (probably in the Golgi; DeMatteis et al. 1993) and is not known to have a role in mitochondrial function. Possible explanations for the Arf1–β-gal localization pattern include (1) Arf1 functions in the mitochondria, (2) Arf1 functions only in secretion and the region containing the fusion protein contains secretory organelles harboring the Arf1–β-gal fusion protein as well as mitochondria, and (3) aggregation of the Arf1-β-gal fusion protein causes clumping of subcellular organelles by an unknown mechanism.

Some of the fusion proteins that localize to discrete sites are derived from previously sequenced ORFs whose functions are unknown. These include the fusions in transformants V76, V145, V8, V160, V319, and V60 (see Table 2; Discussion). Three of these ORFs were identified in the course of systematic sequencing efforts, and the remainder are derived from other sources.

Two fusions correspond to very short ORFs

Surprisingly, in two transformants carrying fusion genes expressed during vegetative growth, *lacZ* is fused to sequences that are present in the GenBank data base, but the fusions are not within long ORFs. One fusion defines an ORF that contains an ATG initiator and is capable of encoding a protein 21 amino acids in length. This ORF lies downstream of the coding region for the single-stranded DNA-binding protein, Ssb1 (Jong et al. 1987), and the fusion protein localizes to the nucleus. Either this ORF encodes a small polypeptide, or this region represents a small exon whose RNA is spliced to upstream sequences, perhaps the *SSB1*-coding region. The small ORF appears to be expressed in vivo, as strains carrying

Table 2. *Vegetative genes. Summary of sequence analysis and disruption phenotypes*

Strain no.	Gene	Spots (no.)	Function	Codon/ total	Disruption phenotype	Comment/ reference	Accession no.
			A. Known genes/ORFs–nuclear				
V9	*REP1*		2μ plasmid stability	124/353	none detected	Hartley and Donelson (1980)	T17581
V10	*H2A1*		histone H2A	76/130	very slow growth; Caf^S; slightly salt-sensitive	Choe et al. (1982)	T17554
V16	*BDF1*[a]		bromodomain protein	22/393	none detected	Z18944	T17560
V130	*BDF1*		bromodomain protein	226/393	not tested	Z18944	T17632
V3	*ADA2*		transcriptional activation	330/434	not tested	Berger et al. (1992)	T17569
V76	*YBR2016*		ORF on chromosome II	362/445	not tested	Holmstroem et al. (1994)	T17577
V170	*ORFZ*		small ORF downstream of *SSB1* gene	20/21[b]	very slow growth	Jong et al. (1987)	T17633
			B. Known genes/ORFs–nucleolar				
V145	*YCL59*		ORF on chromosome III	265/326	not tested	Oliver et al. (1992)	T17558
			C. Known genes/ORFs–mitochondrial				
V83	*ATP1*		mitochondrial F1–ATPase α subunit	447/544	no growth on glycerol	Takeda et al. (1986)	T17579
V87	*ATP1*		mitochondrial F1–ATPase α subunit	376/544	slow growth; no growth on glycerol; slightly Caf^S	Takeda et al. (1986)	T17580
V231	*ATP1*		mitochondrial F1–ATPase α subunit	243/544	not tested	Takeda et al. (1986)	T17562
V338	*NUC1*		mitochondrial nuclease	192/329	no growth on glycerol; orange colonies	Vincent et al. (1988)	T17572
V15	*HEM15*		mitochondrial ferrochelatase	22/393	microcolonies	Labbe-Bois (1990)	T17559
V7	*ARO3*		phospho-2-keto-3 deoxyheptonate aldolase	110/370	none detected	Paravicini et al. (1988)	T17576
V8	*ORFX*		partial ORF upstream of *PPH3*	>141/>298	slow growth on glycerol; slightly Caf^S; ts	Ronne et al. (1991)	T17578
V30	*RNA12*		pre-rRNA maturation	626/850	none detected	Liang et al. (1992)	T17570
			D. Known genes/ORFs–clustered mitochondria				
V5	*ARF1*		ADP–ribosylation factor	31/181	none detected	Sewell and Kahn (1988)	T17574
			E. Known genes/ORFs–cytoplasmic punctate				
V112	*SEC21*	4–8	intra-Golgi transport	345/935	inviable; germination failure	Hosobuchi et al. (1992)	T17555
V120	*TYA*	50	Ty transposon gag	162/440	not tested	Boeke et al. (1988)	T17556
V253	*TYB*	50	Ty transposon protease	13/1328	not tested	Boeke et al. (1988)	T17565
V160[c]	*ORFW*	50	ORF next to *BRF1*	12/>157	inviable; unbudded arrest	Colbert and Hahn (1992)	T17561
V360	*GAC1*	50	possible phosphatase regulatory subunit affecting glycogen accumulation	407/794	none detected	Francois et al. (1992)	T17635
V232	*SEC61*	25	insertion of proteins into the ER	392/480	not tested	Stirling et al. (1992)	T17564
V296	*P450–14DM*	50	cytochrome P450	145/530	inviable; unbudded arrest	Ishida et al. (1988)	T17568
V319	*VKL162*	50–100	ORF on chromosome XI	1268/1483	none detected	Pascolo et al. (1992)	T17571
V354	*RTG2*	50	retrograde regulation gene	122/394	not tested	Liao and Butow (1993)	T17573
			F. Known genes/ORFs–cell periphery				
V236	*CAP/SRV1*	1–4	associated with adenylate cyclase	231/526	not tested	Fedor-Chaiken et al. (1990); Field et al. (1990)	T17563

Table 2. (*Continued*)

Strain no.	Gene	Spots (no.)	Function	Codon/ total	Disruption phenotype	Comment/ reference	Accession no.
			G. Known genes/ORFs–diffuse cytoplasmic				
V60	*ORFY*		DNA from human cDNA library contaminated with yeast sequences	[d]	none detected	Z15419	T17575
V288	*RNR2*		small subunit of ribonucleotide reductase	222/399	inviable; large budded arrest	Elledge and Davis (1987)	T17567
V260	*EXN1*		short ORF lacking ATG near *LEU2–ATE1*		not tested	Chen et al. (1991)	T17566
V142	*PUT4*		proline-specific permease	176/627	none detected	Vandenbol et al. (1989)	T17557
			H. New genes/ORFs–nuclear				
V129					inviable; unbudded arrest		T17586
V25					none detected	homology to Sem5 sex muscle Ab. protein; 72%/54 aa; $P = 3.9 \times 10^{-15}$ (Clark et al. 1992)	T17547
V53					not tested	homology to Nucpl; 84%/13 aa; ($P = 1.5 \times 10^{-4}$ (Vincent et al. 1988)	T17624
V239					none detected		T17545
V34, V42 V43, V124, V138, V152, V163, V228					not tested		T17610, T17620, T17621, T17585, T17588, T17592, T17593, T17600
V41					60% spore lethality		T17619
V227					inviable (or very slow growth); mostly unbudded		T17599
			I. New genes/ORFs–mitochondrial				
V11					not tested		T17583
V86					none detected		T17629
			J. New genes/ORFs–cytoplasmic punctate				
V4		1–5			none detected		T17617
V13		2–6			none detected		T17587
V17		5–10			none detected	homology to Yb103 protein S25330; 61%/63 aa; $P = 8.6 \times 10^{-10}$	T17595
V50, V78		1			none detected		T17622, T17628
V287[e]		1–2			inviable; unbudded arrest		T17551
V363		10–25			none detected	CEN-linked	T17611
V242, V266		1–5			not tested		T17546, T17548
V304, V306, V289		5			not tested		T17602, T17603, T17552
V165		10			very slow growth		T17594
V285		10			none detected		T17550
V364		4–5 (globs)			none detected	homology to yeast Ala/Arg amino-peptidase; 70%/10 aa; $P = 2.9 \times 10^{-27}$ (Caprioglio et al. 1993)	T17612
V6		10–50			none detected	linked to *TRP1*	T17626
V317		1–50			none detected		T17606
V93		10–25			none detected		T17630

(*Table 2 continued on following page*)

Table 2. (*Continued*)

Strain no.	Gene	Spots (no.)	Function	Codon/ total	Disruption phenotype	Comment/ reference	Accession no.
V31		25			none detected	CEN-linked; homology to SYGP ORF 17 L10830; 84%/19 aa; $P = 6.1 \times 10^{-3}$	T17604
V147		25			not tested		T17590
V148		25–50			none detected		T17591
V1, V146, V229, V237, V313, V320		50			none detected		T17582, T17589, T17601, T17544 T17605, T17608
V40		10–50			slow growth; ts; Caf^S; slightly salt-sensitive		T17618
V367		10–50			Caf^S		T17613
V55[f]		50			inviable; large budded arrest		T17625
V123, V321		50			not tested		T17584, T17609
V182		50–100			slow growth		T17596
V183		50			none detected	weakly CEN-linked	T17597
			K. New genes/ORFs–diffuse cytoplasmic				
V22					none detected	CEN-linked	T17598
V32, V369					none detected		T17607, T17615
V27, V291, V368					not tested		T17549, T17553 T17614
V71					Caf^S		T17627
			L. New genes/ORFs–spindle pole body				
V94					inviable; no cdc phenotype		T17631
			M. New genes/ORFs–peripheral patch				
V51					microcolonies		T17623
			N. New genes/ORFs–nuclear rim				
V370					slow growth		T17616

The column labeled Codon/total indicates the codon at which *lacZ* is inserted and the total number of codons in the gene. For genes with homology to known sequences, the percentage of amino acid residues that are identical and the number of amino acids over which the homology extends are indicated. (*P*) The probability that the match is due to chance. For known genes whose sequences are not published, GenBank accession numbers are included in the column labeled Comment/reference. The last column indicates the accession numbers at GenBank for *lacZ*-adjacent sequences. (ts) Temperature-sensitive; (Caf^S) caffeine-sensitive; (ER) endoplasmic reticulum; (aa) amino acids.

[a]Some cytoplasmic staining as well in V16.

[b]Size of ORF based on upstream ATG codon; could be an exon fused to *SSB1*.

[c]Some cytoplasmic dots as well.

[d]ORF size unknown.

[e]Staining near tubulin.

[f]One bright dot; ~50 less intense.

a disruption form very small colonies at 25°C. Another fusion corresponds to a short ORF (called *EXN1* in Table 2) that lacks an ATG initiation codon; sequences expressed from this ORF might also be spliced to upstream coding sequences. In summary, these data indicate that at least some small ORFs in yeast are expressed in vivo.

Disruption of vegetative genes and analysis of mutant phenotypes

The original transformants carrying *lacZ* fusions are heterozygous for the insertion mutations. To determine the phenotypes of haploid cells carrying chromosomal insertion mutations, the diploid transformants were sporulated and tetrads were dissected. Spores from 59 strains carrying fusion proteins that localize to discrete sites, and for which DNA sequence had been obtained, were analyzed for growth at 25°C (Table 2). In nine cases (15%), two spores in each tetrad failed to produce colonies and the two viable spore colonies were Leu^-, indicating that the disrupted gene is essential. In at least six of these transformants, the dead spores germinated and produced progeny cells that arrest at a distinct stage in the cell cycle, suggesting that the disrupted genes correspond to *CDC* genes (Hartwell et al. 1973). From most (50) transformants, four viable spores, two Leu^+ and two Leu^-, were obtained, indicating that the disruption mutations are not lethal. Haploid strains derived from nine of the transformants exhibit moderate or severe growth

defects; in two of these, only microcolonies were visible after 1 week of growth.

The 48 mutant strains that formed sizable Leu$^+$ colonies were incubated under different growth conditions to determine whether they exhibit any detectable growth defects. Spore colonies were replica-plated to rich medium containing (1) glycerol as the primary carbon source, (2) a high concentration of salt, (3) EGTA (a calcium chelator), (4) caffeine (a purine analog), or (5) benomyl (an inhibitor of microtubule assembly). EGTA, caffeine, and benomyl were present at concentrations sublethal to the wild-type strain. Cells were also replica-plated onto rich growth medium and incubated at 37°C, and onto minimal medium (supplemented with appropriate nutrients) and incubated at 25°C.

Under the conditions tested, the majority of insertion mutants grow at the same rate as wild type (summarized in Table 2). None of the mutants are auxotrophic and none are sensitive to EGTA or benomyl. Six mutants are sensitive to caffeine; four fail to grow in the presence of this drug, whereas another two exhibit impaired growth. Two mutants fail to grow on rich medium at 37°C, and two mutants exhibit impaired growth in the presence of 0.9 M NaCl.

Nine transformants carrying fusion proteins that localize to mitochondria were analyzed. Of these, four grow poorly or fail to grow on medium containing glycerol. Three insertions correspond to two genes whose products are known to be important for mitochondrial function (*ATP1*, Takeda et al. 1986; *NUC1*, Liu and Dieckmann 1989); the fourth is encoded by a gene (*ORFX*) that was sequenced previously (Ronne et al. 1991), but was otherwise uncharacterized.

Several transformants display a phenotype in more than one assay (Table 2). For example, transformant V10 exhibits slow grow on rich medium at 25°C, fails to grow in the presence of caffeine, and is slightly sensitive to salt. A total of 15 of the 59 (25%) *lacZ* insertions tested affect nonessential genes but confer defects under some or many conditions of growth.

Identification of genes expressed during meiosis and sporulation

To identify genes expressed during meiosis and sporulation, the yeast transformants screened as described above were replica-plated to sporulation medium and then assayed for β-gal activity on X-gal plates. As controls, cells were also replica-plated to SC − leu, medium limited for nitrogen, and medium limited for glucose. Of 19,000 patches analyzed, 55 strains were identified whose expression was induced significantly after transfer to sporulation plates, but whose β-gal activity was undetectable or barely detectable on control plates. Assuming that there are 6000–8000 genes in the yeast genome (and correcting for insertions into rDNA and for the number of strains containing more than one insertion), these data predict that there are 93–135 meiotically induced genes in yeast.

The DNA sequence adjacent to *lacZ* was determined for 43 independent meiotically induced fusions; these represent a maximum of 40 different genes. Ten fusions correspond to nine genes that have been shown to be induced in meiosis (Table 3A). With two exceptions (*SPS2* and *ISC10*), these genes were identified in genetic screens for mutants defective in meiosis.

Included among the fusion genes whose expression is increased during sporulation are five genes that had been identified previously but were not known to be meiotically induced (*SGV1*, *REV1*, Y′ elements, 25S rRNA, 18S rRNA; Table 3B). In one of these, *REV1* (Larimer et al. 1989), a close match to the consensus URS1 sequence (CCGGCGGCTT in *REV1* vs. PyCGGCGGCTA in the consensus) is located 131 bp downstream of the start of the coding region. URS1 has been shown to be involved in meiosis-specific gene expression (Buckingham et al. 1990). Unexpectedly, four meiotically induced fusions correspond to *lacZ* insertions in sequences that specify rRNAs.

The remaining 26 meiosis-specific insertions do not correspond to genes present in current data bases. However, six of these genes are predicted to encode proteins with amino acid sequence similarity to known proteins (Table 3C). For example, one gene, *MSH4*, encodes a homolog of the bacterial MutS proteins, which are involved in the correction of mismatched base pairs (Modrich 1991). A detailed characterization of this gene will be presented elsewhere (P. Ross-Macdonald and G. Shirleen Roeder, in prep.).

Analysis of meiotically induced fusions by indirect immunofluorescence

Yeast strains carrying all 43 meiotically induced fusions were examined by indirect immunofluorescence using anti-β-gal antibodies; mixed populations of cells representing all stages in meiosis were used for this analysis. These experiments provided information regarding the subcellular location of meiotic gene products as well as the timing of gene expression in meiosis. Of the transformants analyzed, 10 display staining detectable above background.

Several different staining patterns were observed. Of the 10 fusion proteins, 7 display cytoplasmic staining; these include proteins encoded by three known genes, *SPS2*, *HOP1*, and *ZIP1*. The Sps2 and Hop1 fusion proteins exhibit general cytoplasmic staining; the Zip1 fusion protein localizes to a single bright dot per cell, suggesting that the fusion protein forms an aggregate. The Hop1 and Zip1 proteins have been localized previously to meiotic chromosomes (Hollingsworth et al. 1990; Sym et al. 1993); thus, the nuclear localization of these proteins has been disrupted by their fusion to β-gal (see Discussion). Two fusion proteins, Spo11–β-gal and Mek1/Mre4–β-gal, localize to the nucleus (Fig. 4a,b). Mutations in the *SPO11* (Klapholz et al. 1985) and *MEK1/MRE4* (Leem and Ogawa 1992; Rockmill and Roeder 1991) genes confer defects in meiotic recombination and synaptonemal complex formation; *MEK1/*

Table 3. *Meiotically induced genes. Summary of sequence analysis*

Strain no.	Gene	Function	Position	Percent identity/ length (*P*)	Reference	Accession no.
			A. Genes known to be meiotically induced			
M10	*SPS2*	unknown	420/470		Percival-Smith and Segall (1986)	T17506
M41	*SPS2*	unknown	71/470		Percival-Smith and Segall (1986)	T17514
M4	*REC102*	meiotic recombination and chromosome synapsis	170/201		Bhargava et al. (1992); Cool and Malone (1992)	T17503
M54	*HOP1*	component of meiotic chromosomes	346/606		Hollingsworth and Byers (1989); Hollingsworth et al. (1990)	T17523
M46	*MER1*	meiosis-specific RNA splicing	32/270		Engebrecht and Roeder (1990)	T17517
M60	*SPO11*	meiotic recombination and chromosome synapsis	286/398		Klapholz et al. (1985; Atcheson et al. (1987)	T17528
M48	*SPO16*	spore formation	117/199		Malavasic and Elder (1990)	T17518
M69	*MEK1/MRE4*	meiotic recombination and chromosome synapsis	63/497		Rockmill and Roeder (1991); Leem and Ogawa (1992)	T17532
M6	*ISC10*	spore formation	46/267		Kobayashi et al. (1993)	T17504
M85	*ZIP1*	synaptonemal complex structural component	871/875		Sym et al. (1993)	T17543
			B. Genes identified previously but not known to be meiotically induced			
M43	*SGV1*	kinase involved in pheromone-adaptive response	547/657		Irie et al. (1991)	T17515
M57	*REV1*	chemically induced mutagenesis	922/985		Larimer et al. (1989)	T17526
M68	*Y′*	subtelomeric retrotransposon-like element	291/570		Louis and Haber (1992)	T17531
M12	25S rRNA	rRNA	2407 (5)		Georgiev et al. (1981)	T17507
M58	25S rRNA	rRNA	2453 (35)		Georgiev et al. (1981)	T17634
M33	25S rRNA	rRNA	2984 (1)		Georgiev et al. (1981)	T17511
M18	18S rRNA	rRNA	1219 (47)		Rubtsov et al. (1980)	T17509
			C. Novel meiotic fusion genes with products similar to known proteins			
M37	*mutS*	*Salmonella* mismatch repair		20%/522 aa (3.8×10^{-21})	Haber et al. (1988)	T17512
M7	*rab5*	mammalian GTP-binding protein involved in endocytosis		72%/71 aa (7.7×10^{-27})	Chavrier et al. (1990)	T17505
M49	*DHE2*	*Neurospora* NAD-specific glutamate dehydrogenase		34%/77 aa (6.2×10^{-7})	Vierula and Kapoor (1989)	T17519
M81	*GTR*	*Synechocystis* glucose transport protein		35%/86 aa (1.5×10^{-4})	Schmetterer (1990)	T17539
M74	*ATPCU1*	human Cu-transport P-type ATPase		38%/92 aa (5.6×10^{-13})	Vulpe et al. (1993)	T17535
M51	*nodG*	*Rhizobium* dehydrogenase involved in nodulation		26%/78 aa (1.3×10^{-10})	Debelle and Sharma (1986)	T17521

The column labeled Position indicates the codon at which *lacZ* is inserted and the total number of codons in the gene. In the case of insertions in the rDNA, the first number given indicates the distance from the mature 5′ end of the processed rRNA and the number in parentheses indicates the number of codons between the fusion junction and the most proximal ATG codon specified by rDNA. In M12, M58 andM18, the transposon is oriented such that *lacZ* must be transcribed from the strand opposite that of the rRNA. Percent identity/length indicates the percentage of amino acids residues that are identical and the number of amino acids over which the homology extends; (*P*) the probability that the match is due to chance. The last column indicates the GenBank accession number for *lacZ*-adjacent sequences. The accession numbers for transformants M45, M56, and M62, which display defects in meiosis, are T17516, T17525, and T17530, respectively. Accession numbers for other meiotic fusion genes not shown here are T17502, T17508, T17510, T17513, T17520, T17522, T17524, T17527, T17529, T17533, T17534, T17536, T17537, T17538, T17540, T17541, and T17542. (aa) Amino acids.

MRE4 encodes a protein kinase. Neither gene product has been localized previously.

One transformant (M77) displays a novel staining pattern (Fig. 4c). The fusion protein in this strain was detected first as general cytoplasmic staining in a subset of mononucleate cells undergoing the first meiotic division. However, in cells undergoing the second meiotic division, anti-β-gal antibodies stain structures that appear as open circles, with one circle positioned near each pole of each spindle. It is possible that the gene product defined by this insertion is involved in prospore wall formation (Moens 1971; Moens and Rapport 1971). The fusion protein in M77 is not detected in cells that have completed sporulation, indicating that the wild-type protein is not a component of mature spore walls.

Different fusion proteins appeared at different times

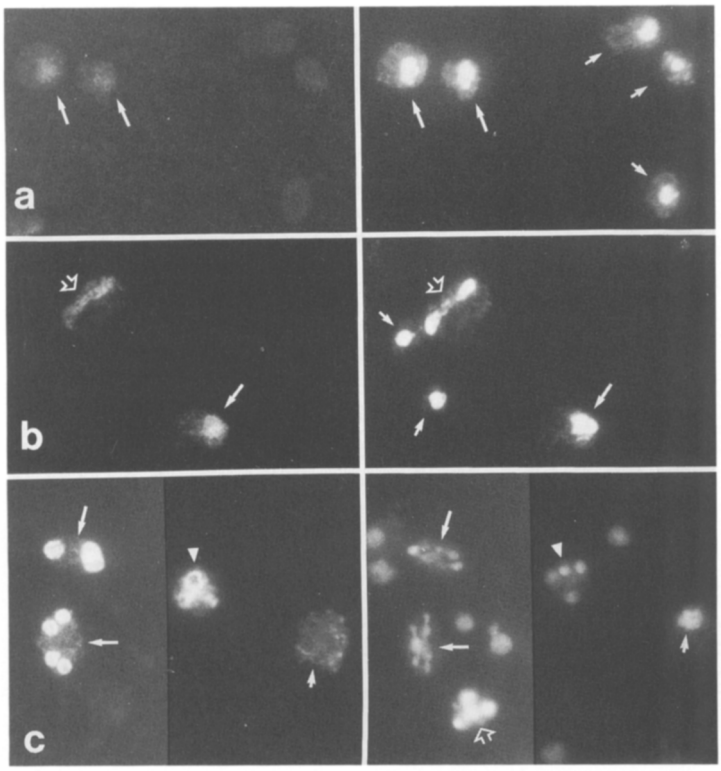

Figure 4. Examples of immunofluorescence patterns observed in sporulating cultures. Staining with anti-β-gal antibodies is shown at *left*; staining of the same cells with the DNA-binding dye, DAPI, is shown at *right*. (*a*) Spo11–β-gal fusion protein (in M69) localizing to the nucleus. Staining is evident in some mononucleate cells (large arrows) but not in others (small arrows). (*b*) Mek1/Mre4–β-gal fusion protein (in M60) localizing to the nucleus. Staining is evident in one cell in which meiosis I is nearly complete (open arrow) and in one mononucleate cell (large arrows); two other mononucleate cells (small arrows) fail to stain. (*c*) Possible prospore wall staining (in M77). In cells in which the second division is ongoing (large arrows) or just completed (arrowhead), anti-β-gal antibodies stain an open circle near each end of each spindle. Staining is not detected in complete asci (open arrow). Most mononucleate cells fail to stain, but rare mononucleate cells, evidently late in prophase I, display general cytoplasmic staining (small arrow) as do binucleate cells (not shown).

during sporulation. The Hop1, Spo11, Mek1/Mre4, and Zip1 fusion proteins were detected in mononucleate cells, consistent with the evidence that these gene products act during meiosis I prophase. The *HOP1* and *MEK1/MRE4* gene products persisted and were detected at later time points as well. The Sps2–β-gal and Isc10–β-gal fusion proteins were detected first in cells undergoing the first division and in cells at later stages in meiosis. In two transformants (M59 and M77), the fusion protein was detected only in binucleate and tetranucleate cells that had completed the first and second meiotic divisions, respectively. In another strain (M83), the fusion protein was detected only in tetranucleate cells. These results are consistent with previous studies demonstrating the existence of four temporally distinct classes of sporulation-specific genes (Clancy et al. 1983; Percival-Smith and Segall 1984; Magee 1987; Law and Segall 1988; Strich et al. 1989; Briza et al. 1990).

Disruption of meiotically induced genes

To determine the phenotypes conferred by insertion mutations in meiotically induced genes, diploids homozygous for each insertion were constructed and assayed for sporulation efficiency and spore viability. In general, the mutants in genes shown previously to be induced in meiosis display phenotypes consistent with the published data, but there were two exceptions. The *zip1* mutant has been shown to arrest in meiosis (Sym et al. 1993), but the *zip1::lacZ* mutant examined in this study sporulates efficiently and produces 50% viable spores. It seems likely that the *zip1::lacZ* fusion analyzed is a null mutation (because the protein fails to localize to the nucleus); thus, it is probable that the difference in phenotype is a consequence of the difference in strain background. Although an *sps2* null mutation has been shown previously to have no effect on meiosis (Percival-Smith and Segall 1986), the *sps2::lacZ* mutant examined in this study displays an approximately threefold reduction in sporulation efficiency.

The mutations in genes identified previously not known to be meiotically induced (Table 3B) did not affect sporulation or spore viability (*SGV1* was not tested). Among the new genes whose products are homologous to known proteins (Table 3C), mutations in two genes confer defects in meiosis. The mutation in strain M7 (Rab5 homolog) reduces sporulation to 10%, and the mutation in M37 (MutS homolog) reduces spore viability to 10%. Of the 20 mutations in genes for which no sequence identities or similarities were detected, three confer meiotic defects. Two mutations (in M45 and M62) reduce spore viability to ~50%, and another (in M56) reduces sporulation efficiency to ~5%.

In summary, *lacZ* insertions in 39 meiotically induced genes were examined for meiotic defects. Of these, a total of 14 confer a reduction in sporulation efficiency and/or spore viability. Of the mutations that confer phenotypes, nine are in genes previously known to be induced in meiosis.

Discussion

Eukaryotic genomes contain many genes of unknown function. The study described here represents the first

systematic approach to characterize the functions of yeast genes on a large scale. To date, we have screened >25% of the yeast genome for genes expressed during vegetative growth and induced during sporulation. Our data indicate that 80% or more of the yeast genome encodes ORFs that are expressed during vegetative growth and that there are 93–135 meiotically induced genes. Our results demonstrate that a significant fraction (at least 3%) of the genes expressed during vegetative growth encode proteins that localize to the nucleus and that a high proportion (at least 1%) of gene products localize to mitochondria.

Some of the genes tagged in the project had been studied already in some detail. Nevertheless, new information about such genes has been generated in this project. For example, the meiosis-specific *SPO11* and *MEK1/MRE4* gene products were not known previously to be localized to the nucleus, and meiotic induction of the subtelomeric Y′ family of repeated DNA sequences had not been demonstrated. In other cases, *lacZ* fusions have been generated in genes that have been sequenced previously but are otherwise uncharacterized; many of these were identified in the course of systematic efforts to sequence the yeast genome. Our project contributes information regarding the expression and function of these ORFS. For example, ORF *YCL59* from chromosome III, ORF *YBR2016* from chromosome II, and ORF *VKL162* from chromosome XI encode proteins that localize to the nucleolus, nucleus, and cytoplasmic dots, respectively. Finally, this project has identified many novel ORFS. The fusion proteins derived from many of these localize to a variety of subcellular locations, including the nucleolus, spindle pole body, bud tips, and prospore walls. We expect that most or all of the new ORFS that we have identified will be sequenced completely within the next few years in the course of systematic sequencing efforts.

Localization of fusion proteins

Fusions to *lacZ* have been used previously to localize a large number of yeast gene products, including proteins that localize to the nucleus (Hall et al. 1984; Silver et al. 1984; Moreland et al. 1985; Moreland et al. 1987; Rutgers et al. 1990; Underwood and Fried 1990), mitochondria (M. Douglas, pers. comm.), spindle pole body (Vallen et al. 1992), microtubules (Meluh and Rose 1990; Page et al. 1994), and sites of polarized growth (Trueheart et al. 1987; C. Costigan and M. Snyder, unpubl.). Nuclear proteins and mitochondrial proteins each contain short localization signals. For mitochondrial proteins, the targeting signals have always been found at or near the amino terminus (Pon and Schatz 1991); similarly, for many nuclear proteins, the localization domain is amino-terminal (Silver 1991). Thus, β-gal fusions are expected to be fairly reliable for localizing proteins to these organelles.

A large number of fusion proteins displaying cytoplasmic staining were identified in our screen. At least in some cases, the localization pattern probably reflects the biology of the protein. For example, the *SEC21* gene is thought to be involved in transport to or within the Golgi apparatus (Hosobuchi et al. 1992), and its product localizes to four to eight cytoplasmic dots, consistent with Golgi staining (Franzusoff et al. 1991; Redding et al. 1991). The number of spots observed with Ty fusions is also consistent with the number of Ty particles expected per cell. Assuming that many of the other cytoplasmic staining patterns reflect correct localization of the authentic proteins (see below), our data are consistent with the concept that the cytoplasm is extremely heterogenous.

Advantages and disadvantages of β-gal fusions

Although β-gal fusions have been used successfully to localize a large number of proteins, there are several concerns associated with their use. First, *lacZ* insertions prevent the expression of coding sequences distal to the insertion site and thereby may remove sequences important for proper localization. For example, in this study, fusion proteins defined by two genes (*HOP1* and *ZIP1*), whose products have been shown to be components of meiotic chromosomes, failed to localize to the nucleus. Hence, this approach probably underestimates the number of proteins that localize to specific subcellular sites (other than the cytoplasm) and overestimates the number of proteins located in the cytoplasm. Second, it is possible that the fusion protein itself may cause localization artifacts. For example, the Arf1–β-gal fusion appears to cause clustering of mitochondria.

In some cases, β-gal fusions have been particularly beneficial. For example, it has not been possible to localize the authentic Kar1 or Kar3 proteins, whereas fusion proteins have been localized (Meluh and Rose 1990; Vallen et al. 1992). In the case of Cik1, β-gal fusions localize to microtubules, whereas the authentic protein targets to the spindle pole body (Page et al. 1994). This observation allowed the identification of a region in the Cik1 protein that interacts with microtubules. Thus, fusion proteins often can provide information concerning the localization or function of a gene product that is not apparent from studies of the bona fide protein. Furthermore, β-gal fusions can aid in the identification of localization sequences. For example, a nuclear localization sequence in the Rep1 protein must lie within the first one-third of the protein (residues 1–123). The Zip1–β-gal fusion protein identified in this study does not localize to the nucleus even though the *ZIP1::lacZ* fusion junction is located only 4 codons from the 3′ end of the coding region. This suggests that the nuclear localization signal is located at the extreme carboxyl terminus of the protein, as predicted from the putative protein sequence (Silver 1991; Sym et al. 1993). In summary, although there are concerns about the use of β-gal fusion proteins, they have been used successfully in a large number of cases and they can prove useful in the identification of localization domains.

Disruption of vegetatively expressed genes

Haploid segregrants carrying insertion mutations were

derived from 59 transformants whose fusion proteins localize to discrete sites (Table 2). The frequency of essential genes (15%) is comparable to that found for random insertions throughout the genome (Goebl and Petes 1986) and for the frequency of disrupted genes analyzed on chromosome I (H. Bussey and D. Kaback, pers. comm.). The frequency of essential genes was not statistically different among transformants displaying different staining patterns (e.g., nuclear staining versus cytoplasmic dots).

The insertion mutations generated by this scheme are not necessarily null mutations. However, for the two genes identified in this study that were known to be essential (*SEC21*, *RNR2*), the insertion mutations rendered cells inviable. Insertion mutations in two known mitochondrial genes (*ATP1* and *NUC1*) rendered the strains unable to grow on glycerol as a carbon source (as expected for mutations in mitochondrial genes) and mutations in genes known to be required for meiosis conferred meiotic defects. Thus, we expect many, if not most, insertions to generate null alleles. Nevertheless, it is likely that some insertions in essential genes do not result in cell inviability; such strains are likely to be hypomorphs and therefore useful for phenotypic analysis.

Analysis of disruption mutations in genes that encode mitochondrial proteins revealed mutants in three genes that fail to grow in glycerol-containing medium; one of these has not been characterized previously. Mutations in several other genes whose fusion proteins localize to the mitochondria did not result in any growth defect on glycerol-containing medium, indicating either that these mutations do not disrupt gene function or that the genes are not essential for respiration. The latter possibility may explain why many genes encoding mitochondrial proteins have not been identified previously even though a large number of genetic screens for mitochondrial mutants have been performed (Pon and Schatz 1991).

Interestingly, a number of the strains identified in this study are sensitive to caffeine. Because caffeine is a purine analog, the high proportion of caffeine-sensitive strains might reflect the fact that ATP or GTP is involved in a high fraction of cellular processes.

Our study indicates that most of the genome encodes proteins that are synthesized during vegetative growth, consistent with previous studies indicating that much of the genome is expressed as RNA (Yoshikawa and Isono 1990). Nevertheless, disruption of most yeast genes does not result in an obvious growth defect. However, by testing a variety of different growth conditions, a large fraction of genes that exhibit a growth defect can be uncovered (25% of total; 30% of nonessential genes). Analysis of additional growth conditions may increase the percentage of mutant strains exhibiting a defect.

For some nonessential genes, functionally redundant homologs may exist elsewhere in the genome. The two genes encoding histone H2A provide one example. Disruption of one of the two genes results in viable cells with a reduced growth rate (this study; Kolodrubetz et al. 1982). However, even for "redundant" genes, it may be possible to identify conditions under which one gene member is essential for growth. In the case of histone H2A, strains containing a disruption of the gene encoding H2A1 fail to grow on medium containing caffeine.

Meiosis-induced genes

The screen for meiotically induced fusion genes is expected to be less sensitive than the screen for vegetative fusions for two reasons: (1) Only a fraction of the cells (~30%) undergo meiosis on sporulation plates; and (2) fewer cells are present on sporulation medium than on growth medium. Due to this reduction in sensitivity, the calculation of 93–135 meiotically induced genes may be an underestimate. However, it should be noted that this screen did identify the *SPO11* and *MER1* genes, which are expressed at very low levels (Atcheson et al. 1987; Engebrecht and Roeder 1990), suggesting that the screen is sufficiently sensitive to identify nearly all meiotically induced genes.

A significant fraction (36%, 14/39) of the meiosis-specific genes identified in this screen confer a defect in meiosis when disrupted. This is in contrast to previous screens for meiotically induced genes, in which the vast majority of genes identified were found to confer no meiotic phenotype when disrupted (Kaback and Feldberg 1985; Percival-Smith and Segall 1986; Law and Segall 1988; Briza et al. 1990; Bishop et al. 1992). Perhaps our ability to identify genes expressed at very low levels accounts for this difference. Previous screens, based on differences in RNA levels between vegetative and meiotic cells, focused on abundant mRNAs. On the basis of the frequency of mutants identified in our screen, we estimate that 30–50 meiotically induced genes are required for wild-type levels of sporulation and/or spore viability. Of the 14 mutations identified that confer a meiotic phenotype, 7 are in genes identified previously in genetic screens for meiotic mutants. Of the five mutations in previously unidentified genes, three cause modest reductions in sporulation efficiency or spore viability and probably would be overlooked in large-scale screens for mutants. Thus, this insertional mutagenesis screen is useful in identifying genes that are important, but not essential, for meiosis.

Most meiotically induced genes (64%) are dispensable for meiosis. As discussed above for vegetative genes, it is possible that functionally redundant homologs exist. Perhaps sporulation under different conditions or in different strain backgrounds would reveal roles for some of the meiosis-specific genes that do not appear to be important for sporulation.

Conclusions and future prospects

Assuming that there are 6000–8000 genes in yeast (Olson 1991) and that each gene is fused to *lacZ* with an equal probability, then analysis of 30,000 transformants carrying productive fusion genes will be required to generate at least one insertion in 98% of yeast genes (and an average of approximately four insertions per gene). These

numbers are probably slight underestimates because not all genes are of equal size, Tn*3* transposition is refractory to certain "cold spots" (Wiater and Grindley 1990), and the frequency with which mutagenized yeast sequences substitute for their genomic copies may vary severalfold among different chromosomal locations. Nevertheless, we expect that this approach can be utilized to identify most yeast genes.

The data that we have generated will be useful to many researchers. Investigators who identify a new gene often will be able to determine immediately its expression pattern and the subcellular location of the fusion protein. Conversely, researchers studying a particular organelle will have access to a bank of insertion mutations in genes whose products localize to that structure. In addition, the method described and the existing library will be useful for many other purposes. The library of plasmids carrying *lacZ* fusions can be used to identify genes whose expression is controlled by a gene of interest, by comparison of β-gal expression in strains that are wild type or mutant for the relevant locus. In addition, the library of yeast transformants can be analyzed for fusion genes expressed at other times in the yeast life cycle, such as during mating or in response to DNA damage. Finally, we expect that this project will be useful in the study of multicellular organisms. Many yeast genes have cognates in other organisms (e.g., Cdc42; Munemitsu et al. 1990; Shinjo et al. 1990), reflecting the fact that many cellular processes are shared between yeast and higher eukaryotes. Consequently, we expect that the information about genes generated by this study will be useful for understanding the role of counterparts identified in humans and other eukaryotes.

Materials and methods

Yeast manipulations

Yeast growth media and standard techniques for the manipulation of yeast have been described by Sherman et al. (1986). The diploid yeast strain (Y800) used to generate most transformants is of the S288C background and has the following genotype:

$$\frac{MATa}{MAT\alpha}\ \frac{leu2\text{-}\Delta98}{leu2\text{-}\Delta98}\ \frac{cry1^{R}}{CRY1}\ \frac{ade2\text{-}101}{ade2\text{-}101}\ \frac{HIS3}{his3\text{-}\Delta200}\ \frac{ura3\text{-}52}{ura3\text{-}52}\ \frac{lys2\text{-}801}{lys2\text{-}801}\ \frac{can1^{R}}{CAN1}\ \frac{trp1\text{-}\Delta1}{TRP1}\ \frac{CYH2}{cyh2^{R}}\ Cir^{0}$$

*Construction and analysis of a library of yeast::*lacZ *insertions*

General cloning procedures were carried out as described by Sambrook et al. (1989). A yeast genomic library was constructed as follows. Yeast DNA was digested to an average size of 3 kbp with *Sau*3A, and the first two nucleotides of the 3′ overhang were filled in with the large fragment of DNA polymerase I and dGTP and dATP. The DNA was size-fractionated in an agarose gel, and the 1.9- to 3.5-kb DNA was ligated into pHSS6-Sal. pHSS6-Sal was derived from pHSS6 (Seifert et al. 1986b) by cleaving with *Not*I and replacing the polylinker with the double-stranded oligonucleotide

5′-GGCCGCGTCGACGC - 3′
3′ - CGCAGCTGCGGGCC-5′

The resulting vector contains a *Sal*I site immediately flanked by *Not*I sites. pHSS6-Sal was cleaved with *Sal*I, and the first two nucleotides were filled in with the large fragment of DNA polymerase I and dCTP and dTTP. The modified genomic *Sau*3A DNA fragments were ligated into the modified vector (Wahl et al. 1987) and transformed into the *E. coli* strain XL1-Blue. Fifty-four random clones were analyzed; all contained inserts, and the average size was 3 kbp. The final library is maintained in 18 separate pools and contains a total of 10^5 recombinants, representing 20 genome equivalents of yeast DNA.

The library pools of *E. coli* were mutagenized separately with a mini-Tn*3*::*lacZ*::*LEU2* using modifications of the procedures of Seifert et al. (1986b). Incubations for bacterial conjugations were carried out for short periods (0.5–1.5 hr) to ensure independent insertions. Approximately 3×10^4–1×10^6 independent transpositions were generated for each pool. DNA was prepared separately from each pool, cleaved with *Not*I, and transformed into yeast by electroporation (Becker and Guarente 1991) or the lithium acetate procedure (Ito et al. 1983). The frequency of strains containing double inserts was similar using both procedures.

Transformants were patched at a density of 100 per plate and screened for β-gal production. Patches of cells, rather than colonies, were tested to achieve greater sensitivity. Fusion genes expressed during vegetative growth were identified by growing cells on SC − leu plates and testing for β-gal activity using the chloroform lysis procedure (Xie et al. 1993). Meiotically induced fusions were identified by replica-plating cells to sporulation medium, synthetic complete medium lacking ammonium sulfate, and complete medium lacking glucose and then screening for β-gal activity using the same procedure. Transformants that turned blue after incubation on sporulation medium, but not on medium limited for nitrogen or glucose, were assumed to carry meiotically induced fusion genes. All transformants that turned blue during vegetative growth or sporulation were retested for β-gal activity in the appropriate assays.

The fraction of the yeast genome expressed during vegetative growth was calculated as follows: Of 20,250 transformants, a total of 22,639 independent insertions were screened—20,250 plus the number that had two insertions (0.118×20,250). One-sixth of the insertions should be in the proper orientation and reading frame (3773). The fraction of the genome expressed is 74.2%, which equals the number of productive transformants observed (2800) divided by the number expected if the entire genome is expressed (3773). The estimated size of the yeast genome, exclusive of rDNA, is 12.49 Mbp, and the amount of rDNA is estimated at 1–2 Mbp (Olson 1991). Therefore, the fraction of the genome that is rDNA is estimated to be 7.4% (1/13.49) to 13.8% (2/14.49) and the non-rDNA fraction ranges from 92.6 to 86.2%. Therefore, the fraction of insertions into non-rDNA sequences is 80.1% (74.2/92.6) to 86.1% (74.2/86.2).

Immunofluorescence

Indirect immunofluorescence was carried out using modifications of published procedures (Snyder and Davis 1988; Snyder 1989; Pringle et al. 1991). For analysis of vegetative fusions, yeast strains were inoculated into 100 μl of YPD in microtiter wells using a pronged transfer device and grown for ~12 hr at 30°C on a vortex shaker. All subsequent incubations and washes were performed in the microtiter dishes. Microtiter dishes were centrifuged for 5 min at 500*g* to pellet cells, as required. Cells were fixed by adding 10 μl of 37% formaldehyde and incubating

at room temperature for 60 min. After fixation, cells were washed three times with solution A (1.2 M sorbitol, 50 mM KPO_4 at pH 7.0), and resuspended in 100 µl of solution A containing 0.1% β-mercaptoethanol, 0.02% glusulase, and 5 µg/ml of Zymolyase 100,000T. After incubation at 37°C for 60 min, cells were washed twice with PBS (150 mM NaCl, 50 mM $NaPO_4$ at pH 7.4) and once with PBS containing 0.1% NP-40. A 1:1000 dilution of a rabbit anti-β-gal antibody (Cappel Laboratories) in PBS plus 2% BSA (bovine serum albumin, fraction V; Sigma) was then added. Previously, anti-yeast antibodies were removed from the anti-β-gal serum by successive incubations of the diluted serum with a 1:1 mixture of Y800 spheroplasts and whole cells (fixed with formaldehyde and treated with NP-40) until background staining was no longer apparent. A rat monoclonal antibody against yeast tubulin (YOL1/34; Kilmartin et al. 1982; Sera-lab, 1:100 final dilution of monoclonal supernatant) was included as a positive control for the staining procedures. After incubation with the primary antibodies overnight at 4°C with shaking, the cells were washed once with PBS, once with PBS plus 0.1% NP-40, and then again with PBS. The cells were then incubated with donkey anti-rat antibodies linked to fluorescein isothiocyanate (Jackson ImmunoResearch Laboratories, 1:100 dilution) and Texas Red-linked anti-rabbit antibodies (Jackson ImmunoResearch Laboratories, 1:200 dilution) for 12 hr and washed as described above, except for an extra wash in PBS plus 0.1% NP-40. Finally, the cells were resuspended in 30 µl of PBS and transferred to individual wells on a polylysine-treated eight-well slide. Forty microliters of mounting solution (70% glycerol, 30% PBS, 2% *n*-propyl gallate, 0.25 µg/ml of Hoechst 33258) was added to each sample.

Asynchronous samples of sporulating cells were obtained by replica-plating patches of cells to sporulation medium and scraping cells off the plate 16–20 hr later. Sporulating cells were processed as described above, except that fixation and spheroplasting were carried out in microcentrifuge tubes, instead of microtiter dishes, and cells were stained with diamidino-2-phenyl-indole (DAPI), instead of Hoechst 33258.

Approximately 30% of the vegetative fusion proteins did not yield significant staining above background. To determine whether proteins localizing to discrete sites were being missed, 74 strains that did not exhibit staining above background were subjected to a sensitive biotin/strepavidin staining protocol that greatly amplifies detection (Page and Snyder 1992). Although many of the resulting samples exhibited general cytoplasmic staining (~30%), no fusion proteins were detected that localize to discrete subcellular locations.

Plasmid rescue and DNA sequence analysis

Yeast strains were transformed individually with 0.5 µg of *Pvu*I-cleaved YIp5 (Botstein et al. 1979), and transformants were selected on medium lacking both leucine and uracil. Yeast genomic DNA was prepared from 2-ml cultures using the DTAB lysis method (Gustincich et al. 1991). Each culture was grown overnight at 30°C and pelleted in an Eppendorf tube. The pellet was washed with 0.5 ml of 1.0 M sorbitol. Cells were spheroplasted at 37°C for 65 min in 250 µl of SEM (1.0 M sorbitol, 50 mM $NaPO_4$ at pH 6.8, 0.1 M EDTA, 0.1% β-mercaptoethanol, 150 µg/ml of Zymolyase 100T). The cells were lysed by the addition of 500 µl of DTAB lysis buffer (8% dodecyltrimethylammonium bromide, 1.5 M NaCl, 0.1 M Tris HCl at pH 8.8, 50 mM EDTA) and incubated at 65°C for 5 min. Genomic DNA was extracted with 500 µl of chloroform, and the aqueous phase was collected after a 2 min spin. Water (750 µl) and of 5% cetyltrimethylammonium bromide (45 µl) in 0.4 M NaCl was added to the aqueous phase. The tube was inverted gently 10 times to mix. After a 5-min spin, the pellet was resuspended in 300 µl of 1.2 M NaCl to exchange the detergent, and 750 µl ethanol was added to precipitate the DNA. The pellet was washed once with 0.5 ml of 95% ethanol, and the DNA was resuspended in 25 µl of TE and 5 µl of RNase A (10 mg/ml). Half of the genomic DNA recovered was digested with *Nsi*I and ligated with T4 DNA ligase in 200 µl of ligation buffer. The ligation was then transformed into *E. coli* by electroporation (Sambrook et al. 1989) or the $MnCl_2$ method (Inoue et al. 1990), and transformants were plated on medium containing ampicillin. Plasmid DNAs were prepared from the resulting colonies and digested with *Eco*RI to verify the presence of the *lacZ*, *LEU2*, and Amp^R genes.

A primer to the 5′ end of the *lacZ* gene (CGTTGTAAAACGACGGGATCCCCCT) was used to sequence the flanking DNA. Sequencing was carried out using the dideoxy chain termination method with the Sequenase version 2.0 kit (U.S. Biochemical). To ensure reproducible results were obtained, two independent YIp5 transformants were sequenced for the following fusions: V5, V13, V24, V30, V105, V138, V152, V160, V163, V182, V185, V253, V260, V266, V287, V288, V289, V306, V313, V319, V338, V370. Sequences were analyzed using the BLAST and FASTA programs (Pearson and Lipman 1988).

DNA gel blot analysis

For DNA gel blot hybridization analysis, genomic yeast DNA was cleaved with *Eco*RI, separated in an 0.8% agarose gel, and blotted overnight onto a nylon membrane (Boehringer Mannheim) according to the directions of the manufacturer. The gel blot was probed with a biotinylated probe corresponding to the 3-kbp *Bam*HI fragment of the *lacZ* gene. Preparation of the probe and gel blot analysis was performed using the Genius 1 DNA Labeling and Detection Kit (Boehringer Mannheim). Analysis of 113 strains producing fusion proteins that localized to discrete sites revealed that 12 (10.6%) contain more than one copy of the *lacZ* transposon. Thus, in most cases, the fluorescence signal observed is due to a unique transposon insertion.

Analysis of disruption phenotypes

To examine the phenotypes of strains containing insertions in genes expressed during vegetative growth, strains were sporulated and 20 tetrads were dissected onto rich (YPD) medium. Spore colonies were replica-plated onto SC − leu, SC − tryptophan, and SC − histidine to confirm that the *LEU2*, *TRP1*, and *HIS3* markers segregate 2:2. Analysis of the *trp1* marker made it possible to determine whether the *LEU2* gene in the transposon was centromere-linked. In instances where two spores formed colonies and two did not, the cells that failed to form colonies were examined under the dissecting microscope prior to replica plating. For those spore colonies in which >85% of the cells were arrested at a particular stage in the cell cyle, the phenotype is indicated in Table 2.

Cells were scored for sensitivity to various growth inhibitors by replica-plating the cells from dissection plates onto rich YPD medium containing (1) 4% glycerol as the primary carbon source, (2) 0.9 M NaCl, (3) 10 mM EGTA, (4) 8 mM caffeine, or (5) 10 µg/ml of benomyl and incubating the plates at 24-25°C. Temperature-sensitive growth was analyzed by growth on YPD plates incubated at 37°C, and growth on minimal medium was determined using yeast minimal medium containing the supplements required by all possible auxotrophic segregants (uracil, lysine, leucine, tryptophan, histidine, and adenine).

To examine the phenotypes of strains containing *lacZ* insertions in meiotically induced genes, diploids homozygous for each insertion mutation were constructed and analyzed. The

original transformants, heterozygous for the insertion, were sporulated, and tetrads were dissected and analyzed as described above. Two haploids of opposite mating type, both carrying the insertion marked by *LEU2*, were mated to each other, and the resulting homozygous diploids were then isolated and sporulated. Sporulation efficiency was monitored in the light microscope by counting at least 100 cells for each strain. Spore viability was assessed by dissecting at least 10 tetrads from each mutant. In the wild-type control, ~30% of the cells sporulated and spore viability was >75%. Diploids carrying insertion mutations were classified as mutant in phenotype if sporulation efficiency was 10% or less and/or if spore viability was 50% or less.

Acknowledgments

We thank Beth Rockmill for helpful comments on the manuscript and Helen Kwan, Alison Beitel, Lara Umansky and Karen Voelkel-Meiman for technical assistance. This work was supported by National Institutes of Health (NIH) grant GM28904 to G.S.R. and NIH grants HG00365 and GM36494 to M.S. N.B. was supported by predoctoral training grant 5T32 GM07499 from the NIH.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Atcheson, C.L., B. DiDomenico, S. Frackman, and R.E. Esposito. 1987. Isolation, DNA sequence, and regulation of a meiosis-specific eukaryotic recombination gene. *Proc. Natl. Acad. Sci.* **84:** 8035–8039.

Becker, D.M. and L. Guarente. 1991. High-efficiency transformation of yeast by electroporation. *Methods Enzymol.* **194:** 182–187.

Berger, S.L., B. Pina, N. Silverman, G.A. Marcus, J. Agapite, J.L. Regier, S.J. Triezenberg, and L. Guarante. 1992. Genetic isolation of *ADA2*: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* **70:** 251–265.

Bhargava, J., J. Engebrecht, and G.S. Roeder. 1992. The *rec102* mutant of yeast is defective in meiotic recombination and chromosome synapsis. *Genetics* **130:** 59–69.

Bishop, D., D. Park, L. Xu, and N. Kleckner. 1992. *DMC1*: a meiosis-specific yeast homolog of *E. coli recA* required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69:** 439–456.

Boeke, J.D. and S.B. Sandmeyer. 1991. Yeast transposable elements. In *The molecular and cellular biology of the yeast* saccharomyces*: Genome dynamics, protein synthesis, and energetics* (ed. J.R. Broach, J.R. Pringle, and E.W. Jones), pp. 193–261. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Boeke, J.D., D. Eichinger, D. Castrillon, and G.R. Fink. 1988. The *Saccharomyces cerevisiae* genome contains functional and nonfunctional copies of transposon Ty1. *Mol. Cell. Biol.* **8:** 1432–1442.

Botstein, D., S.C. Falco, S.E. Stewart, M. Brennan, S. Scherer, D.T. Stinchcomb, K. Struhl, and R.W. Davis. 1979. Sterile host yeasts (SHY): A eukaryotic system of biological containment for recombinant DNA experiments. *Gene* **8:** 17–24.

Briza, P., M. Breitenbach, A. Ellinger, and J. Segall. 1990. Isolation of two developmentally regulated genes involved in spore wall maturation in *Saccharomyces cerevisiae*. *Genes & Dev.* **4:** 1775–1789.

Buckingham, L.E., H.-T. Wang, Z.T. Elder, R.M. McCarroll, M.R. Slater, and R.E. Esposito. 1990. Nucleotide sequence and promoter analysis of *SPO13*, a meiosis-specific gene of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **87:** 9406–9410.

Caprioglio, D.R., C. Padilla, and M. Werner-Washburne. 1993. Isolation and characterization of *AAP1*: A gene encoding an alanine/arginine aminopeptidase in yeast. *J. Biol. Chem.* **268:** 14310–14315.

Casadaban, M. and S.N. Cohen. 1979. Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: *In vivo* probe for transcriptional control sequences. *Proc. Natl. Acad. Sci.* **76:** 4530–4533.

Casadaban, M.J., A. Martinez-Arias, S.D. Shapira, and J. Chow. 1983. β-galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. *Methods Enzymol.* **100:** 293–308.

Chavrier, P., R.G. Parton, H.P. Hauri, K. Simons, and M. Zerial. 1990. Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments. *Cell* **62:** 317–329.

Chen, W.N., E. Balzi, E. Capieaux, M. Choder, and A. Goffeau. 1991. The DNA sequencing of the 17 kb *Hin*dIII fragment spanning the *LEU1* and *ATE1* loci on chromosome VII from *Saccharomyces cerevisiae* reveals the *PDR6* gene, a new member of the genetic network controlling pleiotropic drug resistance. *Yeast* **7:** 287–299.

Choe, J., D. Kolodrubetz, and M. Grunstein. 1982. The two yeast histone genes encode similar protein subtypes. *Proc. Natl. Acad. Sci.* **79:** 1484–1487.

Clancy, M.J., B. Buten-Magee, D.J. Straight, A.L. Kennedy, R.M. Partridge, and P.T. Magee. 1983. Isolation of genes expressed preferentially during sporulation in the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **80:** 3000–3004.

Clark, S.G., M.J. Stern, and H.R. Horvitz. 1992. *C. elegans* cell-signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. *Nature* **356:** 340–344.

Colbert, T. and S.M. Hahn. 1992. A yeast TFIIB-related factor involved in RNA polymerase III transcription. *Genes & Dev.* **6:** 1940–1949.

Cool, M. and R.E. Malone. 1992. Molecular and genetic analysis of the yeast early meiotic recombination genes *REC102* and *REC107/MER2*. *Mol. Cell. Biol.* **12:** 1248–1256.

Debelle, F. and S.B. Sharma. 1986. Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. *Nucleic Acids Res.* **14:** 7453–7472.

DeMatteis, M.A., G. Santini, R.A. Kahn, and A. Luini. 1993. Receptor and protein kinase C-mediated regulation of ARF binding to the Golgi complex. *Nature* **364:** 818–820.

Deschenes, R.J. and J.R. Broach. 1987. Fatty acylation is important but not essential for *Saccharomyces cerevisiae RAS* function. *Mol. Cell. Biol.* **7:** 2344–2351.

Elledge, S.J. and R.W. Davis. 1987. Identification and isolation of the gene encoding the small subunit of ribonucleotide reductase from *Saccharomyces cerevisiae*: DNA damage inducible gene required for mitotic viability. *Mol. Cell. Biol.* **7:** 2783–2793.

Engebrecht, J. and G.S. Roeder. 1990. *MER1*, a yeast gene required for chromosome pairing and genetic recombination, is induced in meiosis. *Mol. Cell. Biol.* **10:** 2379–2389.

Fedor-Chaiken, M., R.J. Deschenes, and J.R. Broach. 1990. *SRV2*, a gene required for *RAS* activation of adenylate cyclase in yeast. *Cell* **61:** 329–340.

Field, J., A. Vojtek, R. Ballester, G. Bolger, J. Colicelli, K. Ferguson, J. Gerst, T. Kataota, T. Michaeli, S. Powers, M. Riggs, L. Rodgers, I. Wieland, B. Wheland, and M. Wigler. 1990. Cloning and characterization of *CAP*, the *S. cerevisiae* gene encoding the 70 kD adenylyl cyclase-associated protein. *Cell* **61:** 319–327.

Francois, J.M., S. Thompson-Jaeger, J. Skroch, and U. Zellenka. 1992. *GAC1* may encode a regulatory subunit for protein phosphatase type I in *Saccharomyces cerevisiae*. *EMBO J.* **11:** 87–96.

Franzusoff, A., K. Redding, J. Crosby, R.S. Fuller, and R. Schekman. 1991. Localization of components involved in protein transport and processing through the yeast golgi apparatus. *J. Cell Biol.* **112:** 27–37.

Garfinkel, D.J., J.D. Boeke, and G.R. Fink. 1985. Ty element transposition: reverse transcriptase and virus-like particles. *Cell* **42:** 507–517.

Georgiev, O.I., N. Nikolaev, A.A. Hadjiolov, K.G. Skryabin, V.M. Zakharyev, and A.A. Bayev. 1981. The structure of the yeast ribosomal RNA genes. 4. Complete sequence of the 25S rRNA gene from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **9:** 6953–6958.

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

Gustincich, S., G. Manfioletti, G.D. Sal, C. Schneider, and P. Carninci. 1991. A fast method for high-quality genomic DNA extraction from whole human blood. *BioTechniques* **11:** 298–302.

Haber, L.T., P.P. Pang, D.I. Sobell, J.A. Mankovitch, and G.C. Walker. 1988. Nucleotide sequence of the *Salmonella typhimurium mutS* gene required for mismatch repair: Homology of MutS and HexA of *Streptococcus pneumoniae*. *J. Bacteriol.* **170:** 197–202.

Hall, M.N., L. Hereford, and I. Hershkovitz. 1984. Targeting of *E. coli* β-galactosidase to the nucleus in yeast. *Cell* **36:** 1057–1065.

Hartley, J.L. and J.E. Donelson. 1980. Nucleotide sequence of the yeast plasmid. *Nature* **286:** 860–865.

Hartwell, L.H., R.K. Mortimer, J. Culotti, and M. Culotti. 1973. Genetic control of the cell division cycle in yeast . V. Genetic analysis of *cdc* mutants. *Genetics* **74:** 267–286.

Hollingsworth, N.M. and B. Byers. 1989. *HOP1*: A yeast meiotic pairing gene. *Genetics* **121:** 445–462.

Hollingsworth, N.M., L. Goetsch, and B. Byers. 1990. The *HOP1* gene encodes a meiosis-specific component of yeast chromosomes. *Cell* **61:** 73–84.

Holmstroem, K., T. Brandt, and T. Kallasoe. 1994. The sequence of a 32,420 bp segment located on the right arm of chromosome II from *Saccharomyces cerevisiae*. *Yeast* (in press).

Hosobuchi, M.M., R. Schekman, and T. Kreis. 1992. *SEC21* is a gene required for ER to golgi protein transport that encodes a subunit of a yeast coatomer. *Nature* **360:** 603–605.

Inoue, H., H. Nojima, and H. Okayama. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96:** 23–28.

Irie, K., S. Nomoto, I. Miyajima, and K. Matsumoto. 1991. *SGV1* encodes a CDC28/cdc2-related kinase required for a Gα subunit-mediated adaptive response. *Cell* **65:** 785–795.

Ishida, N., Y. Aoyama, R. Hata, Y. Oyama, S. Imajo, M. Ishiguro, D.B. Sprinson, and Y. Yoshida. 1988. A single amino acid substitution converts cytochrome P450-14DM to an inactive form, cytochrome P450-SG1: Complete primary structures deduced from cloned cDNAs. *Biochem. Biophys. Res. Comm.* **155:** 317–323.

Ito, H., Y. Fukada, K. Murata and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153:** 163–168.

Jong, A.Y.-S., M.W. Clark, M. Gilbert, A. Oehm, and J.L. Campbell. 1987. *Saccharomyces cerevisiae* SSB1 protein and its relationship to nucleolar RNA-binding proteins. *Mol. Cell. Biol.* **7:** 2947–2955.

Kaback, D.B. and L.R. Feldberg. 1985. *Saccharomyces cerevisiae* exhibits a sporulation-specific temporal pattern of transcript accumulation. *Mol. Cell. Biol.* **5:** 751–761.

Kilmartin, J.V., B. Wright, and C. Milstein. 1982. Rat monoclonal antitubulin antibodies derived by using a new nonsecreting rat cell line. *J. Cell Biol.* **93:** 576–582.

Klapholz, S., C.S. Waddell, and R.E. Esposito. 1985. The role of the *SPO11* gene in meiotic recombination in yeast. *Genetics* **110:** 187–216.

Kobayashi, M., Y. Hotta, and S. Tabata. 1993. Isolation and characterization of a yeast gene that is homologous with a meiosis-specific cDNA from a plant. *Mol. Gen. Genet.* **237:** 225–232.

Kolodrubetz, D., M.C. Rykowski, and M. Grunstein. 1982. Histone H2A subtypes associate interchangeably *in vivo* with histone H2B subtypes. *Proc. Natl. Acad. Sci.* **79:** 7814–7418.

Labbe-Bois, R.L. 1990. The ferrochelatase from *Saccharomyces cerevisiae*: Sequence, disruption, and expression of its structural gene *HEM15*. *J. Biol. Chem.* **265:** 7278–7283.

Larimer, F.W., J.R. Perry, and A.A. Hardigree. 1989. The *REV1* gene of *Saccharomyces cerevisiae*: Isolation, sequence, and functional analysis. *J. Bacteriol.* **171:** 230–237.

Lauer, G.O., T.M. Roberts, and L.C. Klotz. 1977. Determination of the nuclear DNA content of *Saccharomyces cerevisiae* and implications for the organization of DNA in yeast chromosomes. *J. Mol. Biol.* **114:** 507–526.

Law, D.T.S. and J. Segall. 1988. The *SPS100* gene of *Saccharomyces cerevisiae* is activated late in the sporulation process and contributes to spore wall maturation. *Mol. Cell. Biol.* **8:** 912–922.

Leem, S.-H. and H. Ogawa. 1992. The *MRE4* gene encodes a novel protein kinase homologue required for meiotic recombination in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **20:** 449–457.

Liang, S., L. Alksne, J.R. Warner, and F. Lacroute. 1992. *rna12+*, a gene of *Saccharomyces cerevisiae* involved in pre-rRNA maturation. Characterization of a temperature sensitive mutant, cloning and sequencing of the gene. *Mol. Gen. Genet.* **232:** 304–312.

Liao, X. and R.A. Butow. 1993. *RTG1* and *RTG2*: two yeast genes required for a novel path of communication from the mitochondria to the nucleus. *Cell* **72:** 61–71.

Link, A.J. and M.V. Olson. 1991. Physical map of *Saccharomyces cerevisiae* genome at 110-kilobase resolution. *Genetics* **127:** 681–698.

Liu, Y. and C.L. Dieckmann. 1989. Overproduction of yeast viruslike particles by strains deficient in a mitochondrial nuclease. *Mol. Cell. Biol.* **9:** 3323–3331.

Louis, E.J. and J.E. Haber. 1992. The structure and evolution of subtelomeric Y′ repeats in *Saccharomyces cerevisiae*. *Genetics* **131:** 559–574.

Magee, P.T. 1987. Transcription during meiosis. In *Meiosis* (ed. P.B. Moens), pp. 355–382. Academic Press, Orlando, Florida.

Malavasic, M.J. and R.T. Elder. 1990. Complementary transcripts from two genes necessary for normal meiosis in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10:** 2809–2819.

Meluh, P.B. and M.D. Rose. 1990. *KAR3*, a kinesin-related gene required for yeast nuclear fusion. *Cell* **60:** 1029–1041.

Mitts, M.R., D.B. Grant, and W. Heideman. 1990. Adenylate

cyclase in *Saccharomyces cerevisiae* is a peripheral membrane protein. *Mol. Cell. Biol.* **10:** 3873–3883.

Modrich, P. 1991. Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.* **25:** 229–253.

Moens, P.B. 1971. Fine structure of ascospore development in the yeast *Saccharomyces cerevisiae*. *Can. J. Microbiol.* **17:** 507–510.

Moens, P.B. and E. Rapport. 1971. Spindles, spindle plaques, and meiosis in the yeast *Saccharomyces cerevisiae* (Hansen). *J. Cell. Biol.* **50:** 344–361.

Moreland, R.B., H.G. Nam, L.M. Hereford, and H.M. Fried. 1985. Identification of a nuclear localization of a yeast ribosomal protein. *Proc. Natl. Acad. Sci.* **82:** 6561–6565.

Moreland, R.B., G.L. Langevin, R.H. Singer, R.L. Garcea, and L.M. Hereford. 1987. Amino acid sequences that determine the nuclear localization of yeast histone 2B. *Mol. Cell. Biol.* **7:** 4048–4057.

Munemitsu, S., M.A. Innis, R. Clark, F. McCormick, A. Ullrich, and P. Polakis. 1990. Molecular cloning and expression of a G25K cDNA, the human homolog of the yeast cell cycle gene *CDC42*. *Mol. Cell. Biol.* **10:** 5977–5982.

O'Kane, C.J. and W.J. Gehring. 1987. Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci.* **84:** 9123–9127.

Oliver, S.G. et al. 1992. The complete DNA sequence of yeast chromosome III. *Nature* **357:** 38–46.

Olson, M.V. 1991. Genome structure and organization in *Saccharomyces cerevisiae*: Genome dynamics, protein synthesis and energetics. In *The molecular and cellular biology of the yeast* Saccharomyces (ed. J.R. Broach, J.R. Pringle, and E.W. Jones), pp. 1–39. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Page, B. and M. Snyder. 1992. CIK1: A developmentally regulated spindle pole body-associated protein important for microtubule functions in *Saccharomyces cerevisiae*. *Genes & Dev.* **6:** 1414–1429.

Page, B.D., L.L. Satterwhite, M.D. Rose, and M. Snyder. 1994. Localization of the Kar3 kinesin heavy chain-related protein requires the Cik1 interacting protein. *J. Cell Biol.* **124:** 507–519.

Paravicini, G., G. Braus, and R. Huetter. 1988. Structure of the *ARO3* gene of *Saccharomyces cerevisiae*. *Mol. & Gen. Genet.* **214:** 165–169.

Pascolo, S., M. Ghazvini, J. Boyer, L. Colleaux, A. Thierry, and B. Dujon. 1992. The sequence of a 9.3 kb segment located on the left arm of the yeast chromosome XI reveals five open reading frames including the *CCE1* gene and the putative products related to *MYO2* and to the ribosomal protein L10. *Yeast* **8:** 987–995.

Pearson, W.R. and D.J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.* **85:** 2444–2448.

Percival-Smith, A. and J. Segall. 1984. Isolation of DNA sequences preferentially expressed during sporulation in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **4:** 142–150.

———. 1986. Characterization and mutational analysis of a cluster of three genes expressed preferentially during sporulation of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **6:** 2443–2451.

Pon, L. and G. Schatz. 1991. Biogenesis of yeast mitochondria. In *The molecular and cellular biology of the yeast* Saccharomyces cerevisiae: *Genome dynamics, protein synthesis and energetics* (ed. J.R. Broach, J.R. Pringle, and E.W. Jones), pp. 333–406. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Pringle, J.R., A.E. Adams, D.G. Drubin, and B.K. Haarer. 1991. Immunofluorescence methods for yeast. *Methods Enzymol.* **194:** 565–602.

Redding, K., C. Holcomb, and R.S. Fuller. 1991. Immunolocalization of Kex2 protease identifies a putative late Golgi compartment in the yeast *Saccharomyces cerevisiae*. *J. Cell. Biol.* **113:** 527–538.

Rockmill, B.M. and G.S. Roeder. 1991. A meiosis-specific protein kinase homolog required for chromosome synapsis and recombination. *Genes & Dev.* **5:** 2392–2404.

Ronne, H., M. Carlberg, G.Z. Hu, and J.O. Nehlin. 1991. Protein phosphatase 2A in *Saccharomyces cerevisiae*: Effects in cell growth and morphogenesis. *Mol. Cell. Biol.* **11:** 4876–4884.

Rothstein, R. 1991. Targeting, disruption, replacement and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol.* **194:** 281–301.

Rubtsov, P.M., M.M. Musakhanov, V.M. Zakharyev, A.S. Krayev, K.G. Skryabin, and A.A. Bayev. 1980. The structure of the yeast ribosomal RNA genes. I. The complete nucleotide sequence of the 18S ribosomal RNA gene from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **8:** 5779–5794.

Rutgers, C.A., P.J. Sharp, J.V. Riet, C.L. Woldringh, and H.A. Raué. 1990. *In vivo* and *in vitro* analysis of structure-function relationships in ribosomal protein L25 from *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **1050:** 74–79.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schmetterer, G.R. 1990. Sequence conservation among the glucose transporter from the cyanobacterium *Synechocystis* sp. PCC 6803 and mammalian glucose transporters. *Plant Mol. Biol.* **14:** 697–706.

Seifert, H.S., E.Y. Chen, M. So, and F. Heffron. 1986a. Shuttle mutagenesis: A method of transposon mutagenesis for *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **83:** 735–739.

Seifert, H.S., M. So, and F. Heffron. 1986b. Shuttle mutagenesis: A method of introducing transposons into transformable organisms. In *Genetic engineering: Principles and methods* (ed. J.K. Setlow and A. Hollaender), pp. 123–134. Plenum Press, New York.

Sewell, J.L. and R.A. Kahn. 1988. Sequences of the bovine and yeast ADP-ribosylation factor and comparison to other GTP-binding proteins. *Proc. Natl. Acad. Sci.* **85:** 4620–4624.

Sherman, F., G.R. Fink, and J.B. Hicks. 1986. *Methods in yeast genetics: A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Shinjo, K., J.G. Koland, M.J. Hart, V. Narasimhan, D.I. Johnson, T. Evans, and R.A. Cerione. 1990. Molecular cloning of the gene for the human placental GTP-binding protein, G_p (G25K): Identification of this GTP-binding protein as the human homolog of the yeast cell-division cycle protein, CDC42. *Proc. Natl. Acad. Sci.* **87:** 9853–9857.

Silver, P.A. 1991. How proteins enter the nucleus. *Cell* **64:** 489–497.

Silver, P.A., L.P. Keegan, and M. Ptashne. 1984. Amino terminus of the yeast *GAL4* gene product is sufficient for nuclear localization. *Proc. Natl. Acad. Sci.* **81:** 5951–5955.

Snyder, M. 1989. The *SPA2* protein of yeast localizes to sites of cell growth. *J. Cell. Biol.* **108:** 1419–1429.

Snyder, M. and R.W. Davis. 1988. *SPA1*: A gene important for chromosome segregation and other mitotic functions in *S. cerevisiae*. *Cell* **54:** 743–754.

Stirling, C.J., J. Rothblatt, R. Deshaies, M. Hosobuchi, and R. Schekman. 1992. Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. *Mol. Biol. Cell.* **3:** 129–142.

Strich, R., M.R. Slater, and R.E. Esposito. 1989. Identification of

negative regulatory genes that govern the expression of early meiotic genes in yeast. *Proc. Natl. Acad. Sci.* **86:** 10018–10022.

Sym, M., J. Engebrecht, and G.S. Roeder. 1993. ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell* **72:** 365–378.

Takeda, M., W.-J. Chen, J. Saltzgaber, and M.G. Douglas. 1986. Nuclear genes encoding the yeast mitochondrial ATPase complex: Analysis of *ATP1* coding the F1-ATPase alpha subunit and its assembly. *J. Biol. Chem.* **261:** 15126–15133.

Trueheart, J., J.D. Boeke, and G.R. Fink. 1987. Two genes required for cell fusion during yeast conjugation: evidence for a pheromone-induced surface protein. *Mol. Cell. Biol.* **7:** 2316–2328.

Underwood, M.R. and H.M. Fried. 1990. Characterization of nuclear localizing sequences derived from yeast ribosomal protein L29. *EMBO J.* **9:** 91–99.

Vallen, E.A., T.Y. Scherson, T. Roberts, and M.D. Rose. 1992. Asymmetric segregation of the yeast spindle pole body. *Cell* **69:** 505–515.

Vandenbol, M., J.-C. Jauniaux, and M. Grenson. 1989. Nucleotide sequence of the *Saccharomyces cerevisiae PUT4* proline-permease-encoding gene: Similarities between CAN1, HIP1, and PUT4 permeases. *Gene* **83:** 153–159.

Vierula, P.J. and M. Kapoor. 1989. NAD-specific glutamate dehydrogenase of *Neurospora crassa*: cDNA cloning and gene expression during derepression. *J. Biol. Chem.* **264:** 1108–1114.

Vincent, R.D., T.J. Hofmann, and H.P. Zassenhaus. 1988. Sequence and expression of *NUC1*, the gene encoding the mitochondrial nuclease in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **16:** 3297–3312.

Vulpe, C., B. Levinson, S. Whitney, S. Packman, and J. Gitschier. 1993. Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. *Nature Genet.* **3:** 7–13.

Wahl, G.M., K.A. Lewis, J.C. Ruiz, B. Rothenberg, J. Zhao, and G.A. Evans. 1987. Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. *Proc. Natl. Acad. Sci.* **84:** 2160–2164.

Wiater, L.A. and N.D.F. Grindley. 1990. Integration host factor increases the transpositional immunity conferred by γδ ends. *J. Bacteriol.* **172:** 4951–4958.

Willumsen, B.M., A.G. Papageorge, N. Hubbert, E. Bekesi, T. Robins, M. Johnsen, W.C. Vass, and D.R. Lowry. 1986. Mutational analysis of a *ras* catalytic domain. *Mol. Cell. Biol.* **6:** 2646–2654.

Woolford, J.D. 1989. Nuclear pre-mRNA splicing in yeast. *Yeast* **5:** 439–457.

Wu, L.-C.C., P.A. Fisher, and J.R. Broach. 1987. A yeast plasmid partitioning protein is a karoskeletal component. *J. Biol. Chem.* **262:** 883–891.

Xie, K., E.J. Lambie, and M. Snyder. 1993. Nuclear dot antigens may specify transcriptional domains in the nucleus. *Mol. Cell. Biol.* **13:** 6170–6179.

Yoshikawa, A. and K. Isono. 1990. Chromosome III of *Saccharomyces cerevisiae*: An ordered clone bank, a detailed restriction map and analysis of transcripts suggest the presence of 160 genes. *Yeast* **6:** 383–401.

Large-scale analysis of gene expression, protein localization, and gene disruption in Saccharomyces cerevisiae.

N Burns, B Grimwade, P B Ross-Macdonald, et al.

Genes Dev. 1994, **8:**
Access the most recent version at doi:10.1101/gad.8.9.1087