J. Mol. Biol. (1992) **225**, 53–65

Molecular Analysis of *Saccharomyces cerevisiae* Chromosome I

On the Number of Genes and the Identification of Essential Genes Using Temperature-sensitive-lethal Mutations

Steven D. Harris†, Judy Cheng‡, Tom A. Pugh§ and John R. Pringle||

Department of Biology
The University of Michigan
Ann Arbor, MI 48109-1048, U.S.A.

(Received 4 October 1991; accepted 2 January 1992)

Previous analyses of *Saccharomyces cerevisiae* chromosome I have suggested that the majority (>75%) of single-copy essential genes on this chromosome are difficult or impossible to identify using temperature-sensitive (Ts^-) lethal mutations. To investigate whether this situation reflects intrinsic difficulties in generating temperature-sensitive proteins or constraints on mutagenesis in yeast, we subjected three cloned essential genes from chromosome I to mutagenesis in an *Escherichia coli* mutator strain and screened for Ts^- lethal mutations in yeast using the "plasmid-shuffle" technique. We failed to obtain Ts^- lethal mutations in two of the genes (*FUN12* and *FUN20*), while the third gene yielded such mutations, but only at a low frequency. DNA sequence analysis of these mutant alleles and of the corresponding wild-type region revealed that each mutation was a single substitution not in the previously identified gene *FUN19*, but in the adjacent, newly identified essential gene *FUN53*. *FUN19* itself proved to be non-essential. These results suggest that many essential proteins encoded by genes on chromosome I cannot be rendered thermolabile by single mutations. However, the results obtained with *FUN53* suggest that there may also be significant constraints on mutagenesis in yeast. The 5046 base-pair interval sequenced contains the complete *FUN19*, *FUN53* and *FUN20* coding regions, as well as a portion of the adjacent non-essential *FUN21* coding region. In all, 68 to 75% of this interval is open reading frame. None of the four predicted products shows significant homologies to known proteins in the available databases.

Keywords: Saccharomyces cerevisiae; chromosome; mutagenesis; temperature-sensitive mutants; gene-number paradox

1. Introduction

In a wide variety of genetic systems, there are major disparities between the numbers of genes identified using classical genetic analyses and the numbers of genes inferred from analyses of mRNA and polypeptide complexity (for reviews, see Pringle, 1981; Kaback *et al.*, 1984; Pringle, 1987). Chromosome I of the yeast *Saccharomyces cerevisiae* has provided a vivid example of this "gene-number paradox" (Kaback *et al.*, 1984; Diehl & Pringle, 1991; Harris & Pringle, 1991). The size of the chromosome (~250 kb¶: Mortimer *et al.*, 1989; Link & Olson, 1991) and the typical spacing of transcribed regions in yeast (approx. 1/2·2 kb: Coleman *et al.*, 1986; Yoshikawa & Isono, 1990, 1991; Capieaux *et al.*, 1991), suggest that there should be more than 100 transcribed regions on chromosome I. Indeed, approximately 60 discrete transcribed regions have been mapped to the ~160 kb of the chromosome studied to date (Coleman *et al.*, 1986; Steensma *et al.*, 1987; Wickner *et al.*, 1987; Diehl &

† Present address: Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, U.S.A.
‡ Present address: Washington University School of Medicine, St Louis, MO 63110, U.S.A.
§ Present address: Anheuser-Busch Companies, One Busch Place, St Louis, MO 63118-1852, U.S.A.
|| Present address: Department of Biology, University of North Carolina, Chapel Hill, NC 27599-3280, U.S.A.

¶ Abbreviations used: kb, 10^3 bases or base-pairs; EMS, ethyl methanesulfonate; NG, *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine; 5-FOA, 5-fluoroorotic acid; Amp, ampicillin; LB, Luria broth; bp, base-pair(s); ORF, open reading frame.

0022-2836/92/090053-13 $03.00/0

Pringle, 1991; D. Kaback, personal communication; B. Diehl, T. Pugh & J. Pringle, unpublished results). In contrast, only about 15 genes have been identified on chromosome I in classical genetic studies to date (Mortimer *et al.*, 1989).

Kaback *et al.* (1984) used temperature-sensitive (Ts^-) lethal mutations in an attempt to identify all essential genes on chromosome I. A strain monosomic for this chromosome was mutagenized using either ethyl methanesulfonate (EMS) or *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (NG) as mutagen. Thirty-two independent mutations that mapped to chromosome I were isolated; however, these mutations identified only three complementation groups, all of which were previously known. In contrast, molecular analyses of portions of chromosome I have identified at least seven additional, previously unknown, essential genes (Diehl & Pringle, 1991; D. Kaback, personal communication; B. Diehl & J. Pringle, unpublished results); extrapolation to the whole chromosome suggests that it should contain at least 15 to 20 essential genes. It seemed possible that the number of genes identified in the Ts^- mutant hunt had been limited, in part, by the specificity of the mutagens used: both EMS and NG induce predominantly G·C to A·T transitions (which yield only a limited number of amino acid substitutions) and exhibit pronounced hotspots and coldspots for mutagenesis (Prakash & Sherman, 1973; Miller, 1983; Lörincz & Reed, 1986; Kohalmi & Kunz, 1988). However, when the study reported by Kaback *et al.* (1984) was repeated using mutagens (ultraviolet light and nitrogen mustard) of very different specificity (Miller, 1983; Hampsey *et al.*, 1986; Kunz *et al.*, 1987; Kunz & Mis, 1989), all but two of the 19 mutations obtained again fell into the same few genes (Harris & Pringle, 1991).

These results suggest that only a minority of the single-copy essential genes can be identified using Ts^- lethal mutations. To ask if this situation reflects an intrinsic difficulty of generating temperature-sensitive proteins or constraints on mutagenesis in yeast, we subjected three of the cloned essential genes from chromosome I to mutagenesis in an *Escherichia coli* mutator strain (Cox, 1976) and screened for Ts^- lethal mutations in yeast using the plasmid shuffle technique (Boeke *et al.*, 1987). In two cases, we were unable to identify Ts^- lethal alleles. In the third case, we recovered Ts^- lethal alleles, but only at a very low frequency. To determine whether these Ts^- lethal alleles contained one or more than one substitution, we sequenced the four mutant alleles and the corresponding wild-type region. This analysis revealed that the mutations (in

Table 1
Saccharomyces cerevisiae *strains used in this study*

Strain	Relevant genotype[a]	Source/Reference
YS05D	α *leu2 ura3*	H. Fares
BD27-11A	**a** *cdc24-4 leu2 ura3*	B. Diehl
BD2631	**a**/α *cdc24-4*/+ *leu2*/*leu2 trp1*/*trp1 ura3*/*ura3*	Diehl & Pringle (1991)
BD2631T2	**a**/α *fun20*::*URA3*/+ *cdc24-4*/+ *leu2*/*leu2 trp1*/*trp1 ura3*/*ura3*	Diehl & Pringle (1991)
SH804	**a** *fun20*::*ura3 leu2 ura3* [*FUN20 URA3*]	This study[b]
BD2631T3	**a**/α *fun19*::*URA3*::*fun53*/+[c] *leu2*/*leu2 trp1*/*trp1 ura3*/*ura3*	Diehl & Pringle (1991)
JC1	**a** *fun19*::*ura3*::*fun53*[c] *leu2 ura3* [*FUN19 FUN53 URA3*]	This study[d]
SH805	**a** *fun19*::*ura3*::*fun53*[c] *leu2 ura3* [*FUN19 FUN53 LEU2*]	This study[e]
JC3	α *fun12*::*IS10R leu2 ura3* [*FUN12 URA3*]	This study[f]
JC4	α *cdc24-101 leu2 ura3* [*CDC24 URA3*]	This study[g]

[a] Some strains have nutritional markers other than those listed here.

[b] BD2631T2 was transformed to Leu^+ with plasmid pPH1 (see the text), and a meiotic segregant harboring both the gene disruption (Ura^+) and the plasmid (Leu^+) was isolated by tetrad dissection and spread onto SC+5-FOA plates. A resistant clone was selected and transformed with plasmid pPH2. Following non-selective growth, a mitotic segregant that had lost pPH1 was identified and designated SH804. The genotype of SH804 was confirmed by crossing it to YS05D, selecting a Ura^- diploid on SC+5-FOA plates, and dissecting tetrads: all asci segregated 2 : 2 for viability, demonstrating that SH804 maintained the deletion of *FUN20*.

[c] DEL3, which was constructed to inactivate *FUN19* (Diehl & Pringle, 1991), is now known to inactivate *FUN53* as well (see Results).

[d] BD2631T3 was transformed to Leu^+ with plasmid pBDC-L (see the text), and a meiotic segregant harboring both the gene disruption (Ura^+) and the plasmid (Leu^+) was isolated by tetrad dissection and spread onto SC+5-FOA plates. A resistant clone was selected and transformed with plasmid pBDC-U. Following non-selective growth, a mitotic segregant that had lost pBDC-L was identified and designated JC1. The genotype of JC1 (in particular, its retention of the *fun19*/*fun53* deletion) was confirmed as described for SH804 (note [b]).

[e] JC1 was transformed to Leu^+ with plasmid pBDC-L, and a subclone that had lost pBDC-U was selected on SC+5-FOA plates and designated SH805.

[f] The construction of a diploid strain heterozygous for a deletion of *FUN12* is described in the text. This strain was transformed to Leu^+ with plasmid pBDC-L, and a Leu^+, Ura^+ meiotic segregant was obtained by tetrad dissection and streaked onto SC+5-FOA plates to isolate resistant subclones. Such subclones arise by recombination between the *IS10R* repeats flanking the *URA3* gene that marks the deletion (Alani *et al.*, 1987). A stable Ura^- clone was then transformed with plasmid pBDC-U and, following non-selective growth, a mitotic segregant that had lost pBDC-L was isolated and designated JC3. The genotype of JC3 (in particular, its retention of the *fun12* deletion) was confirmed as described for SH804 (note [b]), except that strain BD27-11A was used instead of YS05D.

[g] A strain harboring an unconditional lethal mutation in *CDC24* (*cdc24-101*; A. Bender, personal communication) complemented by plasmid pPB170 (see the text) was transformed to Ura^+ with YEp103(CDC24). Following non-selective growth, a mitotic segregant that had lost pPB170 (Leu^-) was isolated and designated JC4. The genotype of JC4 (in particular, its retention of the *cdc24-101* mutation) was confirmed as described for JC3 (note [f]).

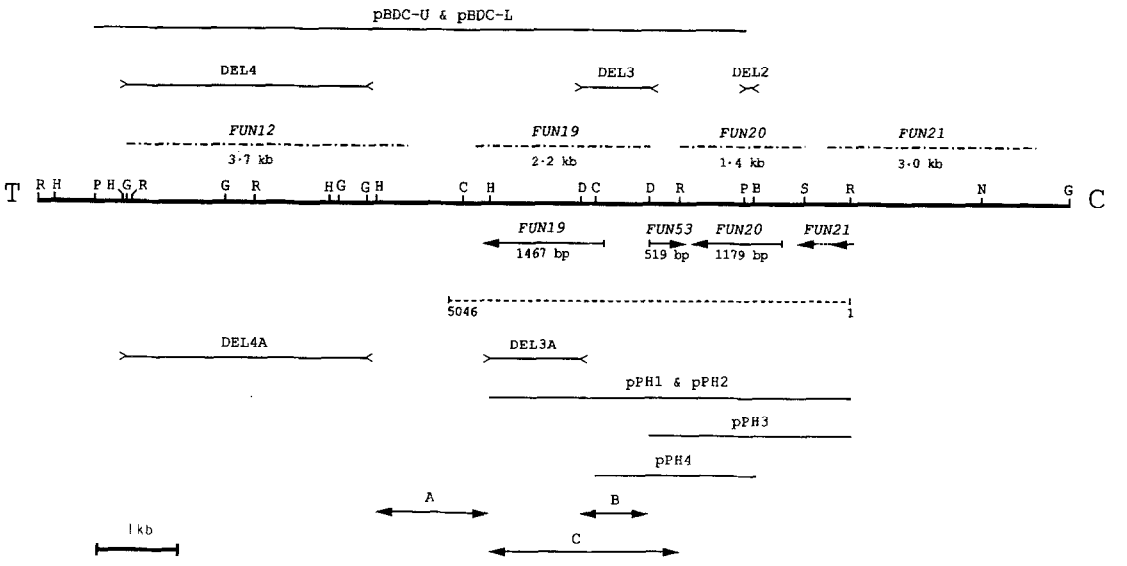

Figure 1. Physical map of the interval studied in the left arm of chromosome I. Plasmids pBDC-U and pBDC-L and the deletions DEL4, DEL3 and DEL2 have been described (Diehl & Pringle, 1991). Shown immediately above the restriction map are the approximate locations, approximate sizes and names of the transcribed regions identified in this interval by RNA–DNA blot-hybridization analyses (Diehl & Pringle, 1991). The large T and C refer, respectively, to the telomere-proximal and centromere-proximal ends of the restriction map. Shown immediately below the restriction map are the approximate locations, sizes and names of the open reading frames identified by sequence analysis of the 5046 bp interval indicated by the broken line (see the text). Given the uncertainty as to which of 2 open reading frames corresponds to *FUN21* (see the text), both are represented (one by the broken line). The deletions DEL4A and DEL3A are described in the text. pPH1, pPH2, pPH3 and pPH4 represent the inserts in the plasmids of these names (see the text). Fragments A and B were used to construct DEL3A. Fragments B and C were used as probes in DNA–DNA blot hybridization experiments (see the text and Fig. 2). Restriction endonuclease cleavage sites are indicated: B, *Bam*HI; C, *Cla*I; D, *Dra*I; G, *Blg*II; H, *Hin*dIII; N, *Nhe*I; P, *Pvu*II; R, *Eco*RI; S, *Sal*I. There may be additional *Dra*I and *Nhe*I sites.

each case, a single substitution) were not in the previously identified *FUN19* gene but in the adjacent, newly identified essential gene *FUN53*. Indeed, *FUN19* appears to be non-essential; the lethality of a deletion constructed to inactivate *FUN19* (Diehl & Pringle, 1991) is in fact due to its effect on *FUN53*.

2. Materials and Methods

(a) *Strains and plasmids*

E. coli strains XL-1 (Bullock *et al.*, 1987) and DH5αf′ (BRL Life Technologies, Gaithersburg, MD) were used for routine plasmid propagation and preparation of single-stranded DNA. *E. coli* strains KC8 (*leuB600 hisB463 pyrF* :: Tn*5 trpC980 strA galUK lacX74*) and W3110 (*mutD5 zaf13* :: Tn*10*) were provided by T. Stearns and E. Cox, respectively. The principal *S. cerevisiae* strains used in this study are described in Table 1. The chromosomal deletions DEL2 (inactivating *FUN20*) and DEL3 (inactivating *FUN53* and presumably *FUN19*; see Results) were both marked with the *URA3* gene (Diehl & Pringle, 1991). In order to use the plasmid shuffle to screen for Ts$^-$ lethal mutations in these genes, it was necessary to inactivate the *URA3* markers while retaining the deletions. This was accomplished by selection of Ura$^-$ subclones on SC+5-FOA plates (Boeke *et al.*, 1987), as described in the notes to Table 1.

Plasmids YEp24, YCp50, pRS315, YRp101(CEN5), YEp103, pNKY150 and YEp352 have been described (Botstein *et al.*, 1979; Rose *et al.*, 1987; Sikorski & Hicter, 1989; Diehl & Pringle, 1991; Coleman *et al.*, 1986; Alani *et al.*, 1987; Hill *et al.*, 1986); their selectable markers are *URA3*, *URA3*, *LEU2*, *LEU2*, *URA3*, *URA3* and *URA3*, respectively. The pBS plasmids were obtained from Stratagene (La Jolla, CA). Plasmids pBDC-U and pBDC-L contain an 8·2 kb *Pvu*II fragment including the *FUN12*, *FUN19* and (as shown here) *FUN53* genes in plasmids YEp24 and YRp101(CEN5), respectively (Diehl & Pringle, 1991). Plasmids YEp103(CDC24) and pBR170 contain a 4·0 kb *Hin*dIII fragment including *CDC24* (Coleman *et al.*, 1986) in plasmids YEp103 and pPB166, respectively (Diehl & Pringle, 1991; A. Bender, personal communication). Plasmid pPB166 is a YEp13 derivative containing the *ADE3* gene (A. Bender, personal communication).

Plasmid pPH1 (Fig. 1) was constructed in 3 steps. First, the 2·3 kb *Hin*dIII–*Eco*RI fragment containing most of *FUN19* and most of *FUN53* was subcloned from pBDC-L into pBS-KS$^-$ to form pA-KS$^-$. Second, the adjacent 2·2 kb *Eco*RI fragment containing part of *FUN53* and all of *FUN20* was inserted into the unique *Eco*RI site of pA-KS$^-$ in the orientation that restored the normal chromosomal configuration of these fragments relative to each other, thus forming pAB-KS$^-$. Finally, the 4·5 kb *Hin*dIII–*Sma*I fragment (the *Sma*I site is from the multiple cloning site of pBS-KS$^-$) was subcloned from pAB-KS$^-$ into *Hin*dIII/*Sma*I-cut pRS315 to form pPH1. Plasmid pPH2 was constructed by subcloning the same 4·5 kb *Hin*dIII–*Sma*I fragment into *Hin*dIII/*Sma*I-cut YEp352. Plasmid pPH3 (Fig. 1) was constructed by subcloning the 2·6 kb *Dra*I–*Sma*I fragment from pPH1 into the *Pvu*II site of YRp101(CEN5). Plasmid pPH4 (Fig. 1) was constructed by subcloning the 2·1 kb *Cla*I–*Bam*HI fragment from pPH1 into YCp50.

(b) *Media and growth conditions*

Yeast strains were grown routinely on rich medium, either YM-P liquid medium (Lillie & Pringle, 1980) or YPD solid medium (Sherman *et al.*, 1986). Media used for the analysis of auxotrophic markers or for selection of transformants were as described (Sherman *et al.*, 1986). 5-Fluoroorotic acid (5-FOA: from PCR, Inc., Gainesville, FL) was used to select yeast strains lacking a functional *URA3* gene (Boeke *et al.*, 1987). 5-FOA was added to the complete minimal medium SC (Sherman *et al.*, 1986) at a concentration of 1·0 mg/ml by dissolving the powder directly in warm medium prior to pouring the plates. Yeast cultures were grown aerobically at 23 °C, 30 °C, or 37 °C, as appropriate. Sporulation was induced in 1% (w/v) potassium acetate at 23 °C. *E. coli* was grown using standard media and procedures (Maniatis *et al.*, 1982).

(c) *Genetic and recombinant DNA manipulations*

Genetic manipulations and transformations of *E. coli* and yeast were performed using standard procedures (Maniatis *et al.*, 1982; Sherman *et al.*, 1986). Yeast transformants were tested for the mitotic stability of their transformed phenotypes as described (Hsiao & Carbon, 1981).

Except as noted, recombinant DNA manipulations were done using standard procedures (Maniatis *et al.*, 1982). DNA fragments were isolated using GENECLEAN II (Bio101, La Jolla, CA) according to the manufacturer's specifications. Total yeast DNA was isolated essentially as described (Bloom & Carbon, 1982). DNA was labeled with [^{32}P]dATP using the oligolabeling procedure (Hodgson & Fisk, 1987). DNA was sequenced by the dideoxy chain-termination method (Sanger *et al.*, 1977) using Sequenase (US Biochemical, Cleveland, OH) and labeling with [^{35}S]dATP. In sequencing the wild-type DNA, both strands were completely sequenced and all sites used for cloning were overlapped by other cloned segments. Clones for sequencing were generated by using exonuclease III to create nested sets of deletions (Henikoff, 1984; Beltzer *et al.*, 1986). Oligonucleotides were synthesized by the University of Michigan Center for Molecular Genetics Oligonucleotide Synthesis Facility (Ann Arbor, MI). The following oligonucleotides were used as primers for the sequencing of the Ts$^-$ lethal alleles of *FUN53*: 5′-TCACTATCAGATTAATG-3′ (nucleotides 2736 to 2720 as shown in Fig. 2) and 5′-TGCAGTTGAAATACTT-3′ (nucleotides 2465 to 2450 as shown in Fig. 2).

DNA–DNA blot hybridizations were performed according to standard procedures (Maniatis *et al.*, 1982) using 0·7% (w/v) agarose gels and nitrocellulose paper (Schleicher & Schuell, Keene, NH). Hybridizations were performed at 42 °C for ~16 h in a solution containing 50% (v/v) formamide, 1 mM-EDTA, 1 M-NaCl, 50 mM-sodium phosphate (pH 6·5), 1 × Denhardt's solution, and 250 μg of sheared and denatured salmon sperm DNA/ml. The most stringent of several washes was performed at 65 °C in a solution containing 0·1 M-NaCl, 0·1% (w/v) SDS, 1 mM-EDTA, 0·1% (w/v) sodium pyrophosphate, 10 mM-Tris (pH 7·4).

(d) *Gene replacements*

Chromosomal deletion/substitution mutations were produced by fragment-mediated gene replacement (Rothstein, 1983). To construct the new deletion of *FUN12* (Fig. 1, DEL4A), the 3·4 kb *Ava*I fragment containing the *URA3* gene was removed from the YEp24 backbone of pBDC-U and the remaining fragment was religated. The resulting plasmid was digested with *Bgl*II, and the 9·5 kb fragment containing the remaining YEp24 backbone along with a small portion of *FUN12* and *FUN12*-flanking DNA was ligated to a 4·5 kb *Bgl*II–*Bam*HI fragment, isolated from pNKY150 (Alani *et al.*, 1987), that contained *URA3* flanked by *IS10R* repeated sequences. Digestion of the resulting plasmid with *Pvu*II yielded the fragment to be used for yeast transformation.

To construct the new deletion of *FUN19* (Fig. 1, DEL3Δ), the 1·5 kb *Hin*dIII fragment A and the 0·9 kb *Dra*I fragment B (see Fig. 1) were subcloned from

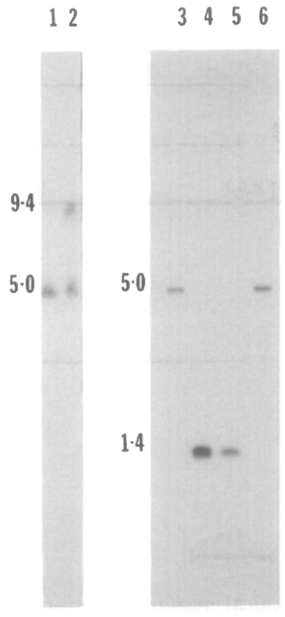

Figure 2. DNA–DNA blot-hybridization analyses of chromosomal DNA from parental and transformed strains. After digestion with the indicated restriction enzymes, DNA fragments were separated and hybridized to radioactively labeled probes as described in Materials and Methods. The sizes of the fragments visualized are indicated in 10^3 bp. Lanes 1 and 2, total DNAs from strain BD2631 (lane 1) and from the same strain following the DEL4A gene replacement (lane 2) was digested with *Eco*RI; the radioactive probe was the 2·3 kb *Hin*dIII–*Eco* RI fragment C (Fig. 1). In the parental strain, a single band of 5·0 kb is seen. The transformant shows this same band (from the copy of chromosome I unaffected by the gene replacement) plus a new band of the predicted size of 9·4 kb (the 3·8 kb *Bgl*II–*Eco*RI segment immediately centromere-proximal to the deletion endpoint plus the 1·1 kb *Eco*RI–*Bgl*II segment immediately centromere-distal to the deletion endpoint plus the 4·5 kb of *IS10R-URA3-IS10R* insert (see Fig. 1)). Lanes 3 to 6, total DNAs from the haploid segregants of 1 complete tetrad from the heterozygous DEL3A diploid were digested with *Eco*RI + *Sma*I; the radioactive probe was the 0·9 kb *Dra*I fragment B (Fig. 1). In the segregants carrying the wild-type configuration of DNA (Ura$^-$, lanes 3 and 6), a single band of 5·0 kb is seen. (There is no *Sma*I site within the 5·0 kb *Eco*RI segment.) The segregants carrying the gene replacement (Ura$^+$, lanes 4 and 5) show a new band of the predicted size of 1·4 kb (the 1·3 kb *Dra*I–*Eco*RI segment (see Fig. 1) plus 60 bp contributed by the *URA3* insert, which has a single *Sma*I site this distance from the chromosomal *Dra*I site).

To construct the new deletion of *FUN19* (Fig. 1, DEL3Δ), the 1·5 kb *Hin*dIII fragment A and the 0·9 kb *Dra*I fragment B (see Fig. 1) were subcloned from

pBDC-L into the *Hin*dIII and *Sma*I sites, respectively, of the pBS-SK$^+$ polylinker in 2 successive steps. That the fragments were in the proper orientation relative to each other was confirmed by verifying that the asymmetrically located *Cla*I site within each fragment was proximal to the *Eco*RI site (between the *Hin*dIII and *Sma*I sites) of the vector. The *URA3* gene, obtained as a 1·1 kb *Hin*dIII fragment from YEp24, was then inserted by blunt-end ligation at the single *Eco*RI site of this plasmid. Digestion of the resulting plasmid with *Cla*I yielded the fragment to be used for yeast transformation.

For each gene replacement, the appropriate fragment was transformed into strain BD2631, selecting for Ura$^+$. In the case of DEL4A, DNA–DNA blot hybridization analysis of the diploid transformant confirmed the correctness of the construction (Fig. 2, lanes 1 and 2), and tetrad analysis showed only 2 viable spores (both Ura$^-$) in each of 12 tetrads. Analysis of DEL3A is described in Results.

(e) *Plasmid mutagenesis*

To mutagenize plasmids pPH3, pBDC-L and pPB170, the *E. coli mutD* mutator strain W3110 was transformed and transformants were selected on M9+Amp plates. For each plasmid, 5 to 20 transformants were selected and pooled. Each pool was grown for 8 to 10 h in LB+Amp and plasmid DNA was isolated. (Because the mutator activity of *mutD* strains is activated in rich media (Cox, 1976), all strain manipulations were carried out on M9 minimal medium except for the specified period of growth in LB.) The level of mutagenesis was checked by testing for loss of *LEU2* function on the mutagenized plasmids. *E. coli* strain KC8 carries a *leuB* mutation that can be complemented by a functional yeast *LEU2* gene (Ratzkin & Carbon, 1977). Therefore, loss of *LEU2* function can be determined by measuring the fraction of the mutagenized plasmids that are no longer able to confer leucine prototrophy to strain KC8. Three independently mutagenized libraries of pBDC-L yielded an average Leu$^-$ frequency of 1·6% (15/943). Two independently mutagenized libraries of pPH3 yielded an average Leu$^-$ frequency of 1·5% (12/814). Two independently mutagenized libraries of pPB170 yielded an average Leu$^-$ frequency of 1·7% (5/296).

(f) *Screen for Ts$^-$ lethal mutations*

The appropriate mutagenized plasmid libraries were introduced into yeast strains SH804, JC1, JC3 and JC4. Transformants were selected on SC-Leu plates at 30 °C and then replica-plated onto duplicate SC+5-FOA and SC-Leu plates that were incubated at 23 °C and 37 °C. Transformants able to grow on SC-Leu at both temperatures and on SC+5-FOA at 23 °C but not at 37 °C were scored as presumptively containing Ts$^-$ lethal mutations in the gene of interest. (The ability of the transformants to grow on SC-Leu at both temperatures verified that the plasmid-borne mutations were in the gene of interest and were not simply affecting the marker gene or some function required for plasmid replication or maintenance.) Transformants able to grow on SC-Leu at both temperatures but unable to grow on SC+5-FOA at either temperature were scored as presumptively containing null mutations in the gene of interest. All presumptive Ts$^-$ mutants, along with a small sample of the putative null mutants, were re-checked by streaking cells from the SC-Leu (23 °C) plate onto SC+5-FOA plates at 23 °C and 37 °C.

Total DNA was prepared from the transformants that presumptively bore Ts$^-$ lethal mutations in *FUN19* or *FUN53*. Plasmids were rescued following passage through *E. coli*, and their structures were confirmed by restriction analysis. These plasmids were then re-transformed into strain JC1 to confirm that the Ts$^-$ lethal phenotype was plasmid-dependent.

(g) *Sequencing of Ts$^-$ lethal mutations*

In preparation for sequencing, the Ts$^-$ lethal alleles of *FUN53* were subcloned into pBS vectors as the 3·2 kb *Hin*dIII–*Pvu*II fragment (Fig. 1). As a control, the same fragment containing the wild-type *FUN53* allele was subcloned into the same vector and sequenced in parallel.

3. Results

(a) *Experimental approach*

From analysis of transcripts by RNA–DNA blot hybridization, Diehl & Pringle (1991) had inferred the existence and locations of *FUN12*, *FUN19*, *FUN20* and *FUN21*, as shown in the upper part of Figure 1. The lethality of deletions DEL2, DEL3 and DEL4, together with the complementation of DEL3, but not DEL2, by pBDC-L, suggested that *FUN20*, *FUN19* and *FUN12* were all essential. To ask why no Ts$^-$ lethal alleles of these genes had been recovered following extensive mutagenesis in yeast (Kaback *et al.*, 1984; Harris & Pringle, 1991), we decided to mutagenize the cloned genes in an *E. coli* mutator strain and then screen for Ts$^-$ lethal mutations using the plasmid shuffle technique (Boeke *et al.*, 1987). Spontaneous mutagenesis is elevated by a factor of 10^4 in *mutD* strains (Cox, 1976), and all types of base substitutions are induced (Fowler *et al.*, 1974; Wu *et al.*, 1990), although the frequencies vary depending on the substrates examined. The plasmid shuffle technique relies upon the use of 5-FOA to select against yeast cells possessing a functional *URA3* gene (Boeke *et al.*, 1987). For each gene, we prepared mutagenized plasmid libraries using a *LEU2*-based plasmid. These libraries were then introduced into a haploid strain carrying a deletion of the gene of interest and a *URA3*-based plasmid carrying the wild-type gene. We then screened for mutants unable to grow on 5-FOA plates at restrictive temperature (37 °C; see Materials and Methods). As a control, we performed parallel studies on *CDC24*, which had yielded numerous Ts$^-$ lethal alleles following mutagenesis in yeast (Kaback *et al.*, 1984; Harris & Pringle, 1991). As described below, in the course of this analysis we discovered that the lethality of DEL3 is due not to its effect on *FUN19* (which proves to be non-essential) but rather to its effect on the adjacent, newly identified gene *FUN53*.

(b) *Isolation of Ts$^-$ lethal mutations*

As expected, *CDC24* readily yielded Ts$^-$ lethal alleles following mutagenesis in the *mutD* strain. Approximately 0·23% of the cells transformed with

Table 2
Summary of screens for Ts⁻ lethal mutants

Gene	No. of transformants screened	No. (%) of Ts⁻ lethal mutants	No. (%) of putative null mutants[a]
CDC24	5090[b]	12 (0·23[c])	131 (2·6[d])
FUN20	21,400[e]	0	304 (1·4[d])
FUN12	27,940[f]	0	264 (0·9[d])
FUN19/FUN53	41,750[g]	4 (0·01[c])	316 (0·8[d])

[a] Samples of the putative null mutants were rechecked by streaking cells from the SC-Leu plate (23 °C) onto SC + 5-FOA plates at 23 °C and 37 °C (see Materials and Methods). These tests showed that only 12 of 37 putative *cdc24* null mutants and 18 of 72 putative *fun20* null mutants were reproducibly negative for growth on SC + 5-FOA at both temperatures. (Presumably, the initial overestimation of the number of null mutants resulted from poor transfer of cells during the initial replica plating.)
[b] Strain JC4 was transformed with mutagenized pPB170 libraries.
[c] The percentage of transformants containing a Ts⁻ lethal mutation.
[d] The percentage of transformants containing a putative null mutation.
[e] Strain SH804 was transformed with mutagenized pPH3 libraries.
[f] Strain JC3 was transformed with mutagenized pBDC-L libraries.
[g] Strain JC1 was transformed with mutagenized pBDC-L libraries.

the mutagenized pPB170 library appeared to contain plasmid-borne Ts⁻ lethal alleles of *CDC24* (Table 2). In addition, 2·6% of the transformants were scored as containing plasmid-borne putative null alleles of *CDC24* (see Materials and Methods), although only approximately 32% of these reproducibly displayed the null phenotype (Table 2, note [a]). Thus, approximately 27% of all mutations isolated in *CDC24* were Ts⁻ lethal mutations.

In contrast, no Ts⁻ lethal allele of either *FUN20* or *FUN12* was obtained despite extensive screening of mutagenized pPH3 and pBDC-L libraries, respectively (Table 2). Putative null alleles also occurred two- to threefold less frequently in *FUN20* and *FUN12* than in *CDC24* (Table 2). In the case of *FUN20* (*FUN12* was not checked), the fraction of putative null mutants that reproducibly displayed the null phenotype was similar to that observed with *CDC24* (Table 2).

In the screen designed to detect *fun19* mutants, plasmid pBDC-L yielded Ts⁻ lethal mutants at a frequency of about 0·01% (Table 2), or approximately 23-fold less frequently than found for *CDC24*. As with *FUN12* and *FUN20*, putative null mutants were also several-fold less frequent than with *CDC24*. If the proportion of putative null mutants that reproducibly displayed the null phenotype was similar to the proportions observed with *CDC24* and *FUN20*, about 4% of all mutations recovered in this plasmid were Ts⁻ lethal mutations.

(c) *Nucleotide sequence of the* FUN19–FUN20 *region*

The complete nucleotide sequence of a 5046 bp region spanning *FUN19* and *FUN20* was determined (Figs 1 and 3). Three complete ATG-initiated open reading frames (ORFs) longer than 103 codons are contained within the region sequenced. The ORF running from nucleotides 912 to 2090 had the position and size expected for the *FUN20* transcribed region identified previously on the basis of a transcript of estimated size 1·4 kb (Figs 1 and 3; and see Diehl & Pringle, 1991). This 393 codon ORF is preceded by an in-frame stop codon at bases −36 to −34 and is followed by two additional in-frame stop codons within 24 bp of the putative primary stop. In addition, two possible polyadenylation sites are located 22 and 124 bp downstream from the putative primary stop. The first ATG upstream from the putative initiation codon is located at −46 to −44 and initiates an ORF of only 17 codons. No TACTAAC consensus splicing signal was observed within or around this ORF. These data suggest that this ORF forms a complete coding region.

The ORF running from nucleotides 3283 to 4749 appears to correspond to the *FUN19* transcribed region (Figs 1 and 3), although it is smaller than expected from the estimated transcript size of 2·2 kb (Diehl & Pringle, 1991). This 489 codon ORF is preceded by an in-frame stop codon at bases −3 to −1 and is followed by an additional in-frame stop codon 94 bp from the putative primary stop. The first ATG upstream from the putative initiation codon is located at −42 to −40 and initiates an ORF of only 13 codons. No TACTAAC consensus splicing signal was observed in or around this ORF. These data suggest that this ORF also forms a complete coding region.

Between these two genes, but encoded by the opposite strand, is an ATG-initiated ORF running from nucleotides 2679 to 2161 (Figs 1 and 3). This 173 codon ORF is preceded by an in-frame stop codon at bases −3 to −1 and is followed by an additional in-frame stop codon 91 bp from the putative primary stop. The first ATG upstream of the putative initiation codon is located at −43 to −41 and is followed immediately by a stop codon. No TACTAAC consensus splicing signal was observed in or around this ORF. These data suggest that this ORF forms a complete coding region, which we

name *FUN53*. This gene probably encodes the approximately 0·6 kb RNA species that was detected using probes from this region by Diehl & Pringle (1991). In the earlier analysis, it was concluded that the approximately 0·6 kb RNA was a breakdown product or an alternative transcript from *FUN19*, primarily on the basis of the argument that there did not seem to be room for another gene in this region given the estimated sizes and positions of the *FUN19* and *FUN20* transcribed regions (Diehl & Pringle, 1991).

In addition to the three complete ORFs, two apparently incomplete, overlapping ORFs (in different reading frames) were detected that terminated within the sequenced region at nucleotides 243 and 634, respectively (Figs 1 and 3). Presumably, one of these ORFs corresponds to the 3′ end of *FUN21*, whose transcript is known to cross the *Eco*RI site that marks the boundary of the sequenced region (Diehl & Pringle, 1991). The shorter of these two ORFs terminates with three consecutive stop codons (at nucleotides 244 to 252), whereas the longer ORF is followed by an in-frame stop codon 12 bp from the putative primary stop. In total, depending on which ORF corresponds to *FUN21*, either 3408 (67·5%) or 3799 (75·3%) of the 5046 nucleotides sequenced are in ORFs.

The putative *FUN19* ORF encodes a polypeptide with a predicted molecular mass of 64,732 Da, a net charge of +51, and six potential asparagine-linked glycosylation sites (Fig. 3). The putative *FUN20* ORF encodes a polypeptide with a predicted molecular mass of 50,380 Da, a net charge of +39, and three potential asparagine-linked glycosylation sites (Fig. 3). The putative *FUN53* ORF encodes a polypeptide with a predicted molecular mass of 22,647 Da, a net charge of +5, and one potential asparagine-linked glycosylation site (Fig. 3). For each gene, hydropathy analysis revealed no obvious candidate for a membrane-spanning domain (data not shown) and the amino terminus did not display the typical features of a signal sequence (Kaiser *et al.*, 1987). Comparison of the predicted primary sequences of the *FUN19*, *FUN20*, *FUN53* and possible *FUN21* polypeptides to other known sequences in the available databases (GenBank, release 67·0; EMBL, release 26·0) revealed no significant homologies.

(d) FUN53 *is an essential gene*

The discovery of an additional gene between *FUN19* and *FUN20* raised two important questions. First, deletion DEL3 removes the first four codons and all of the upstream region of *FUN53* as well as the first 99 codons plus all of the upstream region of *FUN19* (Figs 1 and 3). Thus, the lethality of DEL3 (Diehl & Pringle, 1991) might reflect essentiality of either or both of these genes. Second, since pBDC-L contains complete copies of both *FUN19* and *FUN53*, the associated screen for Ts$^-$ lethal mutants might have identified alleles of either gene.

To determine which gene(s) is (are) essential, we first transformed strain SH805 to Ura$^+$ with plasmid pPH4, which contains all of *FUN53* but only a small part of *FUN19*. Two transformants were then subjected to plasmid stability tests in order to determine if pPH4 could complement the lethality of DEL3. Such complementation would be signalled by the recovery of viable Leu$^-$, Ura$^+$ mitotic segregants, which had lost pBDC-L but retained pPH4. Indeed, 8% (238/3075) of all colonies tested were Leu$^-$, and all of these were Ura$^+$. In contrast, control transformants containing YCp50 instead of pPH4 failed to yield any (933 colonies tested) viable Leu$^-$ mitotic segregants. These data suggest that the lethality of DEL3 is due entirely to the inactivation of *FUN53*.

This conclusion was confirmed by constructing a gene-replacement strain in which a 1043 bp region entirely internal to *FUN19* was replaced by the *URA3* gene (Fig. 1, DEL3A). After transformation of strain BD2631 with the appropriate fragment (see Materials and Methods), three mitotically stable Ura$^+$ transformants were analyzed. For two of these transformants, both tetrad analysis (looking for the expected linkage to *CDC24*) and DNA–DNA blot hybridization analysis indicated that they did not contain the desired gene-replacement event (data not shown). However, analysis of the third transformant by DNA–DNA blot hybridization yielded patterns indicative of the desired gene-replacement event (data not shown). Most of the tetrads dissected from this transformant yielded three or more viable spores (14 with 4 viable spores, 10 with 3, 6 with 2, and 2 with 1), and the *URA3* marker showed an apparent 2 : 2 segregation and the expected linkage to *CDC24* (40 Ura$^+$ Ts$^+$, 41 Ura$^-$ Ts$^-$, 10 Ura$^+$ Ts$^-$, 9 Ura$^-$ Ts$^+$). In addition, DNA–DNA blot hybridization analysis of the four individual segregants from two tetrads indicated that the two Ura$^+$ segregants contained the gene replacement, whereas the two Ura$^-$ segregants had a wild-type configuration of DNA (Fig. 2, lanes 3 to 6, and data not shown).

The results described above suggest that *FUN53* is essential but that *FUN19* is not, and thus that the Ts$^-$ lethal mutations isolated in pBDC-L in the DEL3 strain must be in *FUN53*. To confirm this conclusion, each of the four Ts$^-$ lethal mutants (Ura$^-$, Leu$^+$) was transformed to Ura$^+$ with pPH4 or with YCp50 and three transformants of each type were streaked on SC-Leu plates (to maintain the *LEU2*-based plasmid harboring the mutation) at 37 °C. In each case, as expected, pPH4 was able to complement the Ts$^-$ lethal defect but YCp50 was not.

(e) *Characterization of* FUN53 *Ts$^-$ lethal alleles*

The Ts$^-$ lethal alleles of *FUN53* were sequenced as described in Materials and Methods. Analysis of the sequences revealed that each Ts$^-$ lethal allele contained a single base-pair substitution within

```
GAATTCTACA AGACATCCGT CAATAGCTCC ACCTTCAAAG CTCAACAATC AAAGGTCCAA TTCTTTGCAG TCCTCAACCA TGACGTTAAA TCAAAAAATC GTCCAAGATA  110
CAGTGCGGCA TTTGATGAAC AAAAGTGCAT CTACACCCAA TCCTTCTGCA TCTTCATCGT TAGCGCCTTC ACCAAAAGTA TCCAGTATAA ACAACACCTC CTCAGGAAAA  220
TCATCAAGTA CCCTGATTGC CAATAGTAGT GATACTTTGG CTATAGAGAC ATTAACTCTA GATCCGGAGT CAAACTCCAG TGAATTGTCC ATCAAAAGGG TGAGATTTGC  330
TGGTGTACCG CCAATGACAG AGGCAGAAAA TCCTAAACCA ACAAAGGTGG GTTGGTACAA GAAGCCGGCT GTGCTACATT ATCCGCCAAT ACCTGCTTCG GCAATGATCA  440
AGCCGCTACA GCATAAGTCC AAATATAACA CGCTGAGACA AGAAGAAGGG TTTACCTTTC GAAAAAGTCT TCGAGATGGA TTAGAGTGGG AAAATGGCGA ATCTGGCTCC  550
GAAACTACCA TGATGCCGTT TGGGATCGAA ATCAAAGAGT CGACCGGGCA TAGAATAGCT TCTAAAATCA GAAGCAAATT GAGATAACTC CGCACCTTAT AAAAACAACT  660
CTTTATGTCC CTGGAAAAAA AAAAGAATAC ATATCTAATA CTATTCATAT CCTTCATTGT TCTTGTCTGA TGATTCAAAA GTACTCATAC TATTTTTTAA TATTTAATCT  770
TTTACCTTAA CGTATTATTG TAATTCTTCA CGAATTTGAT TTTTCCCCAG CTAAAAAGTT GATGAAAGAA GTTCAAATTA TGGGGTTATT GACACATGGC AGCTCTGAGC  880
CGAGAGGACG TATCAGCAAC CTCAACCAAA T ATG TTT AGT AAC AGA CTA CCA CCT CCA AAA CAT TCT CAA GGA CGA GTT TCG ACG GCT TTG AGC    974
                                  Met Phe Ser Asn Arg Leu Pro Pro Pro Lys His Ser Gln Gly Arg Val Ser Thr Ala Leu Ser

TCA GAT CGC GTT GAG CCG GCA ATA TTG ACT GAC CAA ATC GCT AAA AAC GTT AAG CTC GAT GAT TTT ATT CCA AAG AGA CAG TCT AAT TTC  1064
Ser Asp Arg Val Glu Pro Ala Ile Leu Thr Asp Gln Ile Ala Lys Asn Val Lys Leu Asp Asp Phe Ile Pro Lys Arg Gln Ser Asn Phe

GAA CTA TCG GTT CCT TTG CCA ACG AAA GCA GAA ATC CAA GAA TGT ACA GCA AGA ACC AAG TCA TAC ATT CAG CGG CTT GTG AAT GCG AAA  1154
Glu Leu Ser Val Pro Leu Pro Thr Lys Ala Glu Ile Gln Glu Cys Thr Ala Arg Thr Lys Ser Tyr Ile Gln Arg Leu Val Asn Ala Lys

CTA GCC AAC TCA AAT AAC AGG GCA TCA TCA AGG TAC GTC ACC GAA ACA CAT CAG GCA CCC GCG AAT CTA TTA TTG AAC AAC AGC CAC CAT  1244
Leu Ala Asn Ser Asn Asn Arg Ala Ser Ser Arg Tyr Val Thr Glu Thr His Gln Ala Pro Ala Asn Leu Leu Leu Asn Asn Ser His His

ATT GAG GTA GTG TCC AAG CAA ATG GAT CCA TTG TTG CCA AGG TTC GTT GGG AAG AAG GCG AGA AAG GTT GTA GCA CCC ACA GAA AAC GAC  1334
Ile Glu Val Val Ser Lys Gln Met Asp Pro Leu Leu Pro Arg Phe Val Gly Lys Lys Ala Arg Lys Val Val Ala Pro Thr Glu Asn Asp

GAA GTC GTG CCT GTT CTC CAT ATG GAT GGC AGC AAT GAT AGG GGA GAA GCT GAT CCA AAT GAG TGG AAG ATA CCT GCA GCT GTG TCA AAC  1424
Glu Val Val Pro Val Leu His Met Asp Gly Ser Asn Asp Arg Gly Glu Ala Asp Pro Asn Glu Trp Lys Ile Pro Ala Ala Val Ser Asn

TGG AAA AAT CCA AAT GGT TAT ACC GTG GCC TTG GAA AGA CGT GTA GGT AAA GCT CTT GAC AAC GAA AAT AAT ACC ATC AAC GAT GGG TTT  1514
Trp Lys Asn Pro Asn Gly Tyr Thr Val Ala Leu Glu Arg Arg Val Gly Lys Ala Leu Asp Asn Glu Asn Asn Thr Ile Asn Asp Gly Phe

ATG AAG CTC TCC GAA GCG TTA GAA AAC GCT GAC AAG AAG GCA AGA CAA GAG ATC AGG TCC AAA ATG GAA TTG AAG CGG CTT GCT ATG GAA  1604
Met Lys Leu Ser Glu Ala Leu Glu Asn Ala Asp Lys Lys Ala Arg Gln Glu Ile Arg Ser Lys Met Glu Leu Lys Arg Leu Ala Met Glu

CAG GAA ATG CTT GCT AAA GAA TCT AAA TTG AAA GAA TTG AGC CAA CGA GCC AGA TAC CAC AAC GGG ACT CCG CAG ACG GGA GCA ATA GTT  1694
Gln Glu Met Leu Ala Lys Glu Ser Lys Leu Lys Glu Leu Ser Gln Arg Ala Arg Tyr His Asn Gly Thr Pro Gln Thr Gly Ala Ile Val

AAG CCC AAA AAG CAA ACG AGC ACA GTG GCC AGA CTA AAA GAG CTG GCG TAC TCT CAA GGA AGA GAC GTA TCC GAA AAG ATA ATT CTG GGC  1784
Lys Pro Lys Lys Gln Thr Ser Thr Val Ala Arg Leu Lys Glu Leu Ala Tyr Ser Gln Gly Arg Asp Val Ser Glu Lys Ile Ile Leu Gly

GCA GCA AAG CGT TCA GAA CAA CCG GAT CTG CAG TAC GAT TCA AGA TTT TTC ACA AGA GGG CAA ATG CCT CCG CCA AAA GGC ATG AAG ACC  1874
Ala Ala Lys Arg Ser Glu Gln Pro Asp Leu Gln Tyr Asp Ser Arg Phe Phe Thr Arg Gly Gln Met Pro Pro Pro Lys Gly Met Lys Thr

AGG TTT ATG ACA ACC CAC TGT TCG TCC AAC AAG ATA TTG AAA GCA TAT ACA AGA CCA ACT ACG AAA AGC TGG ACG AAG CGG TCA ATG TTA  1964
Arg Phe Met Thr Thr His Cys Ser Ser Asn Lys Ile Leu Lys Ala Tyr Thr Arg Pro Thr Thr Lys Ser Trp Thr Lys Arg Ser Met Leu

AGA GTG AAG GTG CCA GTG GTT CTC ACG GCC CCA TTC AGT TTA CTA AAG CTG AAT CCG ATG ATA AAT CGG ATA ACT ATG GCG CCT AGG CCC  2054
Arg Val Lys Val Pro Val Val Leu Thr Ala Pro Phe Ser Leu Leu Lys Leu Asn Pro Met Ile Asn Arg Ile Thr Met Ala Pro Arg Pro

AGG ATG AGC ACT AGG AAA ACA AAG CAT TCT TGT GCT TGA GTT GTA TTA TAG ATA AAA AAA TAA AAC TAC TGC ATA//ATG CAC ATG TAG    2141
Arg Met Ser Thr Arg Lys Thr Lys His Ser Cys Ala ***

                  *** Asp Asp Glu Asn Glu Asn Glu Asn Glu Arg Gly Ile Lys Lys Phe Asp Asn Ile Ile Phe Asp Asn Asp
ATA TAT AAT TAT ATA A GAT TAG CAG GAG CAA AAG CAA AAG CAA AAG GGA AGG TTA AAA GAA CTT TAG TAA CTA TTA TTT CAG TAA TAG    2229

Ser Leu His Ser Ser Gln Ser Cys Lys Ile Ile Asn Leu Ile Lys Ser Asn Arg Arg Met Ala Phe Gln Glu Ile Lys Lys Ile Thr Gly
CGA CTC CAC ACT ACT AAC TGA TGT GAA CTA ATA CAA GTC TTA AAA TCT TAA GGA AGA GTA TCG TTT GAC GAG ATA AAA GAA CTA CCA GGG  2319

Ser Val Lys Val Pro Asn Val Ile Leu Gly Asp Val Asp Gly Ile Lys Ser Met Leu Met Leu Ala Met Ile Val Leu Asp Cys Asp Glu
TGA ATG GAA CTG CCC CAA GTG CTA GTC AGG TAG CTG CAG CGG TTA AAA GCT GTA GTT GTA GTT ACG GTA CTA TTG TTC CAG CGT TAG GAG  2409
```

Figure 3. Nucleotide sequence of the 5046 bp *FUN19-FUN53-FUN20* interval from chromosome I and predicted amino acid sequences of the *FUN19*, *FUN53* and *FUN20* products. The nucleotide sequence is numbered relative to the G of the *Eco*RI site within *FUN21* that marks the centromere-proximal boundary of the sequenced region (Fig. 1). The sequence of the strand in the 5′ to 3′ orientation starting from this *Eco*RI site is shown up to nucleotide 2129 and from nucleotides 2747 to 5046. Between the // marks at positions 2130 and 2746 (encompassing the *FUN53* open reading frame), the sequence of the opposite strand is presented. The nucleotide substitutions (at positions 2417 and 2423) identified in the 4 Ts^- lethal alleles of *FUN53* are shown along with the corresponding amino acid substitutions. The following restriction sites are indicated by underlining: *Eco*RI (1, 2276), *Sal*I (589), *Bam*HI (1268), *Pvu*II (1411), *Dra*I (2666, 3573), *Cla*I (3387, 4952) and *Hin*dIII (4618). The putative primary stop codons are indicated by *** and the possible polyadenylation sites following the *FUN20* open reading frame are doubly underlined. These data have been deposited with GenBank/EMBL, Accession no. X62577.

```
        Tyr     Thr
         ^       ^
Arg His Cys Arg Ile Ile Gly Thr Ser Thr Lys Asn Ser Phe Tyr Lys Leu Gln Leu Leu Ser Asn Cys Lys Ala Ser Gly Tyr Asp Gly
AGC TAC CGT AGC CTA ATA GGG GCA TCT GCA GAA TAA ACT TTT CAT AAA GTT GAC GTT CTC TCT CAA TGT AAA ACG TCT GGG TAT CAG CGG  2499
        A       C

Leu Asn Leu Ser Leu Ser Arg Arg Ile Glu Gln Leu Ile Ser Lys Ile Ser Val Asp Ala Pro Ser Ala Arg His His Ser Leu Leu Ile
GTC TAA GTT GCT GTC GCT CGC AGC ATA GAG AAC CTC ATA GCT GAA ATA CCT GTG TAG GCG TCC GCT ACG AGA CAC TAC GCT TTC GTT CTA  2589

Asp Ala Lys Ser Val Ser Glu Glu Val Asn Thr Asp Thr Pro Pro Phe Ile Ile Glu Phe Leu Ile Tyr Arg Ser Lys Leu Arg Val Met
CAG ACG AAA GCT GTG TCT AAG GAG TTG CAA CCA CAG ACA TCC ACC CTT ATA TTA AAG TTT TTC CTA TAT AGA TGA AAA ATT TGC ATG GTA  2679

AATAGCTGAA CAAAGAAACA ACGTGATTGA TAACGATAGT GTAATTAGAC TATCACTTTT CATTTTA//A TTTTTCAAGA GCGCTAAGCA AGAACAAAAC AAAACAAAGT 2787
CAGGAGGCAA TTAAAGACAA CAAGATAGAA AAACCAGAGT GGTTGCTGTG GTAGTACGTT GGGGGCTCGT TTGTCGAGGT TTTTAAGCAG GTTTGAATGG AAGAAGAAAA 2897
AAACGAGCAC TTACTGGAGT TACTTGGTAA GAAAAGAGTA AGAGGGTTTG GTGGGTGATA TATATTTTCA TGCCACCAAC TGCGAAAAGA CAAGGCATGC ATCCATGGCT 3007
GGTGAAGTGT TTGTTTAGGA TATGAACTTC AACCCAATTC AGTTCCCTCA CAAACATCGT CTTTCTTCCT TCACCCACAG GTGCAGGCGC CGCATGGGTG AGCCTGAGAC 3117
AGGGTCAACA CATTAGGTCG ATTCAAAGTT ACGACAGTCA TACAGATAAT GAAACGACAT ATTTGAACTG GTAAACGAGC TTGCCATTGG CTCTGAGTAT GCCCAATGCT 3227
TCTTTAGTGA ATAATGTTCA CTAAAACTGT TGGTACATCC GTGCACCTCC GCTAG ATG AGG TGC AGG GGT CTA TCA AGA ACC CCC TTC TTC CTA AGA GAA 3327
                                                            Met Arg Cys Arg Gly Leu Ser Arg Thr Pro Phe Phe Leu Arg Glu

CCG CTT ATA AAA GTA GCC ACA TGG TTC ATA ACA TCC CCG TTA TGG CTC ATA CAC TGG GCA TCG ATA AAC AGG ATA TTC GCG CAT CGT AAC  3417
Pro Leu Ile Lys Val Ala Thr Trp Phe Ile Thr Ser Pro Leu Trp Leu Ile His Trp Ala Ser Ile Asn Arg Ile Phe Ala His Arg Asn

ATT TAT AGG GGA AAC ATG GGA TTA TAT TCT CCT GAA TCT GAA AAG TCT CAA TTA AAT ATG AAT TAC ATT GGT AAG GAT GAT TCG CAG TCC  3507
Ile Tyr Arg Gly Asn Met Gly Leu Tyr Ser Pro Glu Ser Glu Lys Ser Gln Leu Asn Met Asn Tyr Ile Gly Lys Asp Asp Ser Gln Ser

ATT TTC AGA CGT CTG AAT CAG AAT TTG AAA GCA AGT AAC AAC AAT AAC GAT AGT AAT AAA AAC GGT TTA AAC ATG AGT GAT TAT AGC AAT  3597
Ile Phe Arg Arg Leu Asn Gln Asn Leu Lys Ala Ser Asn Asn Asn Asn Asp Ser Asn Lys Asn Gly Leu Asn Met Ser Asp Tyr Ser Asn

AAT TCA CCC TAT GGG CGC TCG TAC GAC GTA AGA ATT AAC CAG AAC TCA CAA AAT AAT GGA AAT GGA TGC TTT TCT GGC AGC ATT GAC TCC  3687
Asn Ser Pro Tyr Gly Arg Ser Tyr Asp Val Arg Ile Asn Gln Asn Ser Gln Asn Asn Gly Asn Gly Cys Phe Ser Gly Ser Ile Asp Ser

TTG GTT GAT GAA CAT ATA ATA CCA TCG CCA CCT TTG TCG CCC AAG CTG GAG TCG AAA ATT AGC CAC AAT GGC TCA CCC CGC ATG GCC TCT  3777
Leu Val Asp Glu His Ile Ile Pro Ser Pro Pro Leu Ser Pro Lys Leu Glu Ser Lys Ile Ser His Asn Gly Ser Pro Arg Met Ala Ser

TCA GTG CTA GTG GGA TCT ACG CCT AAA GGC GCT GTA GAG AAT GTG CTG TTC GTG AAG CCT GTA TGG CCC AAT GGG TTA TCA AGA AAA AGG  3867
Ser Val Leu Val Gly Ser Thr Pro Lys Gly Ala Val Glu Asn Val Leu Phe Val Lys Pro Val Trp Pro Asn Gly Leu Ser Arg Lys Arg

TAC CGC TAC GCC ACC TAC GGG TTT CTG TCT CAA TAC AAA ATT TTC AGC AAT TTG GCC CAA CCA TAT TCT AAG AAC ATT ATC AAC CGG TAC  3957
Tyr Arg Tyr Ala Thr Tyr Gly Phe Leu Ser Gln Tyr Lys Ile Phe Ser Asn Leu Ala Gln Pro Tyr Ser Lys Asn Ile Ile Asn Arg Tyr

AAC AAT CTG GCC TAT AAT GCT AGA CAT AAA TAT TCC AAA TAC AAT GAT GAT ATG ACT CCT CCT CCT CTG CCT TCC TCT TCT TCT AGA TTA  4047
Asn Asn Leu Ala Tyr Asn Ala Arg His Lys Tyr Ser Lys Tyr Asn Asp Asp Met Thr Pro Pro Pro Leu Pro Ser Ser Ser Ser Arg Leu

CCT TCC CCG TTA GCA TCT CCG AAT TTG AAT AGA CAA GCA AGA TAT AAT ATG AGG AAA CAG GCT CTC TAC AAT AAC AAT CTA GGA AAG TTT  4137
Pro Ser Pro Leu Ala Ser Pro Asn Leu Asn Arg Gln Ala Arg Tyr Asn Met Arg Lys Gln Ala Leu Tyr Asn Asn Asn Leu Gly Lys Phe

GAA TCC GAC ACT GAA TGG ATA CCA CGG AAA CGC AAG GTA TAC TCA CCA CAA AGA AGA AGC ATG ACT ACC AGT CCA CAT CGC GCC AAG AAG  4227
Glu Ser Asp Thr Glu Trp Ile Pro Arg Lys Arg Lys Val Tyr Ser Pro Gln Arg Arg Ser Met Thr Thr Ser Pro His Arg Ala Lys Lys

TTC TCA CCC TCT GCA TCC ACT CCT CAC ACT AAC ATT GCA TCC ATT GAG GCG ATT CAT GAT GCT CCT CAA TAT ATA CCA AAC GTC TCA TGG  4317
Phe Ser Pro Ser Ala Ser Thr Pro His Thr Asn Ile Ala Ser Ile Glu Ala Ile His Asp Ala Pro Gln Tyr Ile Pro Asn Val Ser Trp

AAA AAA TTA CCA GAT TAC TCT CCG CCC TTA TCT ACG CTT CCT ACA GAC AGT AAC AAG TCA CTC AAG ATC GAG TGG AAG GGG TCT CCA ATG  4407
Lys Lys Leu Pro Asp Tyr Ser Pro Pro Leu Ser Thr Leu Pro Thr Asp Ser Asn Lys Ser Leu Lys Ile Glu Trp Lys Gly Ser Pro Met

GAC CTG TCC ACA GAC CCG CTG AGG AAC GAG CTA CAC CCT GCT GAA CTA GTT CTC GCT CAA ACT CTA AGG TTA CCT TGT GAT TTG TAT CTG  4497
Asp Leu Ser Thr Asp Pro Leu Arg Asn Glu Leu His Pro Ala Glu Leu Val Leu Ala Gln Thr Leu Arg Leu Pro Cys Asp Leu Tyr Leu

GAT TCT AAG AGA AGG TTA TTT TTG GAA AAG TTT ATA GAT CAA AGA AAG GGT TGC CGT TTA GAA GGA CCG ACG CCC AAA AAG CCT GTA GGA  4587
Asp Ser Lys Arg Arg Leu Phe Leu Glu Lys Phe Ile Asp Gln Arg Lys Gly Cys Arg Leu Glu Gly Pro Thr Pro Lys Lys Pro Val Gly

TCG ACG TTA ATA AAG CAT CAA GAC TAT TCC AAG CTT TCG AGA AGG TTG GCT GGC TAC AGG ATT CGA ATT TTA CGA AGT ACT TAT AAT ACA  4677
Ser Thr Leu Ile Lys His Gln Asp Tyr Ser Lys Leu Ser Arg Arg Leu Ala Gly Tyr Arg Ile Arg Ile Leu Arg Ser Thr Tyr Asn Thr

CAT GGT GGT ATG CAT TTT AAA CAT GGT TTT CTT TTA TCC AAT TTT TTA GCA TTT TTC TAC ATA ATA CAT TCG TGA ATA CAT ACA TCA TAT  4767
His Gly Gly Met His Phe Lys His Gly Phe Leu Leu Ser Asn Phe Leu Ala Phe Phe Tyr Ile Ile His Ser ***

CATAGCACAT ACTTTTTCCC ACTTTATATT ATAACACATC ATCAAGTAAA TCCAATTGAG GCTCTTCAAT ATCTATACTA ACTATATCCC CTGCTTGCGT CCTCGCATCC 4877
GGCTTGGCAG ACTGTTTTTC GTCCTCTTTA TCACCTAGTA GCCCCAATTT TCGTAGGACA TTATTCGTCC TTATATCGAT ATATCGTGTT CTGAAATCCT TATTATTTTG 4987
CTTCCCGTTG GTCTTCAATC TCTGCAGTAG TGCCTCTACC TCCTCTGTTG AACTGGTAC                                                        5046
```

Fig. 3.

FUN53. In three cases (*fun53-10*, *fun53-32* and *fun53-54*), an identical G·C to A·T transition had occurred, resulting in the substitution of cysteine 88 with tyrosine (Fig. 3). In the fourth case (*fun53-46*), an A·T to G·C transition had occurred, resulting in the substitution of isoleucine 86 with threonine (Fig. 3). These substitutions must reflect the occurrence of at least three independent mutational events, as *fun53-54* was recovered from a mutagenized library distinct from the one that yielded *fun53-10* and *fun53-32*. In confirmation of the conclusions described above, sequencing of *FUN19* from the four mutant plasmids revealed no substitution in three cases, although in the fourth case (*fun99-10*), the plasmid contained two separate G·C to A·T transitions, resulting in the substitution of alanine 86 with valine and of arginine 303 with lysine.

4. Discussion

Previous studies of *S. cerevisiae* chromosome I have suggested that the majority (>75%) of the single-copy essential genes on this chromosome are difficult or impossible to identify using Ts$^-$ lethal mutations, regardless of the mutagen used (see Introduction; Kaback *et al.*, 1984; Harris & Pringle, 1991). There appear to be two factors that may contribute to this problem. First, it is possible that for the majority of proteins there is no single mutation (or even a pair of mutations) that can produce temperature sensitivity over an appropriate temperature range and hence a Ts$^-$ lethal mutation. Consistent with this view, intensive efforts using *in vitro* mutagenesis failed to identify Ts$^-$ lethal mutations in the yeast tubulin, histone or *RAD3* genes (Han *et al.*, 1987; Huffaker *et al.*, 1988; Schatz *et al.*, 1988; Naumovski & Friedberg, 1987). Moreover, sequencing of multiple, independently isolated alleles of *CDC8* (Su & Sclafani, 1991; R. Sclafani, personal communication) and *CDC28* (Lörincz & Reed, 1986) has suggested that there are very few sites within these genes that can be mutated to yield Ts$^-$ gene products. However, it is also possible that there exist major constraints on mutagenesis and/or repair within certain genes or chromosomal domains in yeast. For example, the seemingly complex nature of the mechanisms controlling mutagenesis (Lawrence, 1982; Friedberg *et al.*, 1991) might result in different regions of chromosome I being differentially susceptible to mutagenesis. Indeed, there is evidence for differential repair in the genomes of yeast, humans, and Chinese hamster ovary cells; in particular, DNA sequences that are transcriptionally active appear to be more efficiently repaired in these organisms (Bohr *et al.*, 1985; Mellon *et al.*, 1987; Leadon & Snowden, 1988; Terleth *et al.*, 1990). There is also evidence for the preferential induction of mutations on the transcribed strand of a yeast tRNA gene (Armstrong & Kunz, 1990).

To begin assessing the relative contributions of these two factors, we attempted to generate Ts$^-$ lethal mutations in three genes that had not been identified during extensive conventional Ts$^-$ mutant hunts (Kaback *et al.*, 1984; Harris & Pringle, 1991) but had been identified and shown to be essential during molecular analyses of chromosome I (Diehl & Pringle, 1991). The approach used to pass plasmids containing these genes through a *mutD* mutator strain of *E. coli* (Cox, 1976), then to screen for Ts$^-$ lethal mutations in yeast using the plasmid shuffle technique (Boeke *et al.*, 1987). As all types of base substitutions are induced at high frequencies in *mutD* strains (Fowler *et al.*, 1974; Wu *et al.*, 1990), this approach should generate a relatively non-biased distribution of mutational events in each gene. Indeed, control experiments showed that Ts$^-$ lethal mutations were obtained easily in *CDC24*, which had also yielded many such mutations during the conventional mutant hunts (Kaback *et al.*, 1984; Harris & Pringle, 1991). Nevertheless, our results suggest that two of the genes tested cannot readily yield Ts$^-$ lethal mutations, whereas the third gene yields Ts$^-$ mutations, but only at low frequency.

We were unable to obtain Ts$^-$ lethal mutations in either *FUN12* or *FUN20*. In such cases, it seems likely that there is no single mutation that can produce a protein that functions at one temperature but is non-functional at another temperature just 10 to 15 deg.C higher. Indeed, studies of both phage λ repressor and phage T4 lysozyme have demonstrated that each of these proteins is remarkably tolerant of a diverse array of amino acid substitutions (Alber *et al.*, 1988; Reidhaar-Olson & Sauer, 1988; Lim & Sauer, 1989; Bowie *et al.*, 1990). Moreover, the structural requirements for many proteins may be satisfied so long as the core is composed of hydrophobic residues while the surface remains hydrophilic (Alber *et al.*, 1987; Lim & Sauer, 1989; Bowie *et al.*, 1990). It would probably be very difficult, if not impossible, to obtain Ts$^-$ lethal mutations in the genes encoding proteins exhibiting this level of structural redundancy. Perhaps complex combinations of amino acid substitutions would yield appropriately thermolabile gene products; however, such events would be rare at the level of mutagenesis we used, and the chances of recovering such mutants following conventional *in vivo* mutagensis would be small.

Although we were able to obtain Ts$^-$ lethal mutations in *FUN53*, the frequency was approximately 23-fold lower than that observed for *CDC24*. This could be due, in part, simply to the smaller size of *FUN53* (ORF of 173 codons *versus* 854 for *CDC24* (Miyamoto *et al.*, 1991)). In this respect, it should be noted that putative null mutations were also about threefold less frequent for *FUN53* than for *CDC24*. In addition, *FUN53* may have fewer appropriately mutable sites per unit length of open reading frame than does *CDC24*: DNA sequence analysis revealed that three of the four Ts$^-$ lethal mutations recovered in *FUN53* were due to identical mutations at the same position, suggesting that there are very few mutational events that will render the *FUN53* product thermolabile over an

appropriate temperature range. (Although genetic allelism tests suggest that Ts⁻ *cdc24* mutations occur at 5 or more sites (M. McAllister, R. Preston & J. Pringle, unpublished results), no precise number is available.) Finally, there may be hotspots for *mutD* mutator activity (Wu *et al.*, 1990) within *CDC24* or coldspots within *FUN53* (at the few relevant positions) that also contribute to the difference in mutability between these genes as observed in the present experiments.

We were surprised to observe that single base substitutions (in particular, G·C to A·T and A·T to G·C transitions) could generate Ts⁻ *fun53* alleles. Such substitutions are readily induced by the mutagens employed during the conventional mutant hunts (Kaback *et al.*, 1984; Harris & Pringle, 1991), yet no Ts⁻ lethal alleles of *FUN53* were recovered (in contrast to 25 Ts⁻ lethal *cdc24* mutations). Although it remains possible that the different yields of Ts⁻ *cdc24* and *fun53* mutants in the conventional mutant hunts reflects simply a difference in the total number of relevant mutable sites (as discussed above), it seems more likely that there are also significant constraints on mutagenesis in yeast. These constraints might be quite local (e.g. mutational coldspots due to unfavorable nucleotide context at the few relevant sites within *FUN53*). However, the fact that *fun53* mutations were recovered neither with EMS or NG (Kaback *et al.*, 1984) nor with ultraviolet light (Harris & Pringle, 1991), mutagens that typically have different distributions of hotspots and coldspots (Miller, 1983; Kunz *et al.*, 1987; Kohalmi & Kunz, 1988), suggests that there may be some broader reduction of mutability affecting the *FUN53* gene itself or the chromosomal region in which it lies.

In the course of this work, we determined the complete nucleotide sequence of the 5046 bp region including *FUN19*, *FUN53* and *FUN20*. Given the uncertainty as to the actual 3′ end of the apparently incomplete *FUN21* ORF, we estimate that 68 to 75% of the region sequenced is open reading frame. This number is consistent with estimates of the density of open reading frames obtained from other large regions of the yeast genome that have been sequenced (Melnick & Sherman, 1990; Thierry *et al.*, 1990; Chen *et al.*, 1991; C. Davies, personal communication; S. Oliver & A. Goffeau, personal communication), as well as with data for the close packing of genes obtained by transcript analysis (Coleman *et al.*, 1986; Yoshikawa & Isono, 1990, 1991; Capieaux *et al.*, 1991). It is striking, but not particularly unusual, that the *FUN20* and *FUN53* ORFs terminate within 70 bp of each other.

At the beginning of this study, we were not aware of the existence of *FUN53*. Its presence was suggested when sequencing of *FUN19* from plasmids ostensibly carrying Ts⁻ alleles of this gene found no substitution in three of four cases. Subsequently, sequencing of the entire region revealed ORFs clearly corresponding to the *FUN19* and *FUN20* transcribed regions as well as an additional ORF located between these two genes. In retrospect, it seems clear that the low abundance 0·6 kb transcript detected previously using probes from this region (Diehl & Pringle, 1991) is encoded by the newly identified gene, *FUN53*. The experiments reported here demonstrate that the lethality of DEL3, previously attributed to its effect on *FUN19*, is in fact due to its effect on *FUN53*, and that *FUN19* itself is non-essential.

These results emphasize the importance of utilizing both DNA sequencing and transcript analysis in attempts to define genomic organization in yeast and other organisms. On the one hand, it is easy to misinterpret faint bands detected by RNA–DNA blot hybridization, particularly in regions that are crowded with genes. Moreover, some genes may not be expressed detectably under the conditions used or in the cell type(s) examined. On the other hand, it can be difficult to discern from sequence analysis alone which ORFs correspond to *bona fide* genes, particularly when the ORFs are short or in systems where genes may have multiple introns.

In summary, the studies reported here have contributed in two ways to understanding of the "gene-number paradox" in yeast. First, the sequence analysis has shown that the packing of genes in this portion of chromosome I is even tighter than earlier studies based on transcript mapping (Diehl & Pringle, 1991) had suggested. Thus, the magnitude of the failure of classical genetic approaches to identify genes on chromosome I, and presumably in the genome as a whole, is probably even greater than previously demonstrated. Second, our results suggest that many of the single-copy essential genes on chromosome I cannot be rendered temperature-sensitive by single or simple combinations of substitutions. Moreover, even some genes that can yield Ts⁻ lethal alleles by single substitutions do so at extremely low frequency. Thus, the continuing heavy reliance upon Ts⁻ lethal mutations for the identification of interesting essential genes means that many genes governing important cellular processes will be overlooked, unless this approach is complemented extensively by alternative genetic strategies.

We thank D. Kaback, B. Diehl, A. Bender, H. Fares, A. Goffeau, S. Oliver, Y. Ohya and Y. Anraku for encouragement, helpful discussions, the communication of unpublished results, strains and plasmids. This study was supported by USPHS (NIH) grant GM31006 and by a predoctoral fellowship to S.H. from the Horace H. Rackham School of Graduate Studies, The University of Michigan.

References

Alani, E., Cao, L. & Kleckner, N. (1987). A method for gene disruption that allows repeated use of *URA3* selection in the construction of multiply disrupted yeast strains. *Genetics*, **116**, 541–545.

Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C. & Matthews, B. W. (1987). Temperature-sensitive mutations of bacteriophage T4 lysozyme occur at sites with low mobility and low solvent accessibility in the folded protein. *Biochemistry*, **26**, 3754–3758.

Alber, T., Bell, J. A., Dao-pin, S., Nicholson, H., Wozniak, J. A., Cook, S. & Matthews, B. W. (1988). Replacements of Pro^{86} in phage T4 lysozyme extend an α-helix but do not alter protein stability. *Science,* **239**, 631–635.

Armstrong, J. D. & Kunz, B. A. (1990). Site and strand specificity of UVB mutagenesis in the *SUP4-o* gene of yeast. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 9005–9009.

Beltzer, J. P., Chang, L.-F. L., Hinkkanen, A. E. & Kohlhaw, G. B. (1986). Structure of yeast *LEU4*. *J. Biol. Chem.* **261**, 5160–5167.

Bloom, K. S. & Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell,* **29**, 305–317.

Boeke, J. D., Truehart, J., Natsoulis, G. & Fink, G. R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* **154**, 164–175.

Bohr, V. A., Smith, C. A., Okumoto, D. S. & Hanawalt, P. C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the *DHFR* gene of CHO cells is much more efficient than in the genome overall. *Cell,* **40**, 359–369.

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

Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. & Sauer, R. T. (1990). Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science,* **247**, 1306–1310.

Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Biotechniques,* **5**, 376–378.

Capieaux, E., Ulaszewski, S., Balzi, E. & Goffeau, A. (1991). Physical, transcriptional and genetical mapping of a 24 kb DNA fragment located between the *PMA1* and *ATE1* loci on chromosome VII from *Saccharomyces cerevisiae*. *Yeast,* **7**, 275–280.

Chen, W., Balzi, E., Capieaux, E., Choder, M. & Goffeau, A. (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.

Coleman, K. G., Steensma, H. Y., Kaback, D. B. & Pringle, J. R. (1986). Molecular cloning of chromosome I DNA from *Saccharomyces cerevisiae*: isolation and characterization of the *CDC24* gene and adjacent regions of the chromosome. *Mol. Cell. Biol.* **6**, 4516–4525.

Cox, E. C. (1976). Bacterial mutator genes and the control of spontaneous mutation. *Annu. Rev. Genet.* **10**, 135–156.

Diehl, B. E. & Pringle, J. R. (1991). Molecular analysis of *Saccharomyces cerevisiae* chromosome I: identification of additional transcribed regions and demonstration that some encode essential functions. *Genetics,* **127**, 287–298.

Fowler, R. G., Degnen, G. E. & Cox, E. C. (1974). Mutational specificity of a conditional *Escherichia coli* mutator, *mutD5*. *Mol. Gen. Genet.* **133**, 179–191.

Friedberg, E. C., Siede, W. & Cooper, A. J. (1991). Cellular responses to DNA damage in yeast. In *The Molecular and Cellular Biology of the Yeast* Saccharomyces: *Genome Dynamics, Protein Synthesis, and Energetics* (Broach, J. R., Pringle, J. R. & Jones, E. W., eds), pp. 147–192, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hampsey, D. M., Das, G. & Sherman, F. (1986). Amino acid replacements in yeast iso-1-cytochrome *c*. *J. Biol. Chem.* **261**, 3259–3271.

Han, M., Chang, M., Kim, U.-J. & Grunstein, M. (1987). Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication and transcription. *Cell,* **48**, 589–597.

Harris, S. D. & Pringle, J. R. (1991). Genetic analysis of *Saccharomyces cerevisiae* chromosome I: on the role of mutagen specificity in delimiting the set of genes identifiable using temperature-sensitive-lethal mutations. *Genetics,* **127**, 279–285.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene,* **28**, 351–359.

Hill, J. E., Myers, A. M., Koerner, T. J. & Tzagoloff, A. (1986). Yeast/*E. coli* shuttle vectors with multiple unique restriction sites. *Yeast,* **2**, 163–167.

Hodgson, C. P. & Fisk, R. Z. (1987). Hybridization probe size control: optimized 'oligolabelling'. *Nucl. Acids Res.* **15**, 6295.

Hsiao, C.-L. & Carbon, J. (1981). Direct selection procedure for the isolation of functional centromeric DNA. *Proc. Nat. Acad. Sci., U.S.A.* **78**, 3760–3764.

Huffaker, T. C., Thomas, J. H. & Botstein, D. (1988). Diverse effects of β-tubulin mutations on microtubule formation and function. *J. Cell. Biol.* **106**, 1997–2010.

Kaback, D. B., Oeller, P. W., Steensma, H. Y., Hirschman, J., Ruezinsky, D., Coleman, K. G. & Pringle, J. R. (1984). Temperature-sensitive lethal mutations on yeast chromosome I appear to define only a small number of genes. *Genetics,* **108**, 67–90.

Kaiser, C. A., Preuss, D., Grisafi, P. & Botstein, D. (1987). Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science,* **235**, 312–317.

Kohalmi, S. E. & Kunz, B. A. (1988). Role of neighbouring bases and assessment of strand specificity in ethylmethanesulphonate and N-methyl-N′-nitro-N-nitrosoguanidine mutagenesis in the *SUP4-o* gene of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **204**, 561–568.

Kunz, B. A. & Mis, J. R. A. (1989). Mutational specificities of 1,3-bis(2-chloroethyl)-1-nitrosourea and nitrogen mustard in the *SUP4-o* gene of *Saccharomyces cerevisiae*. *Cancer Res.* **49**, 279–283.

Kunz, B. A., Pierce, M. K., Mis, J. R. A. & Giroux, C. N. (1987). DNA sequence analysis of the mutational specificity of U.V.-light in the *SUP4-o* gene of yeast. *Mutagenesis,* **2**, 445–453.

Lawrence, C. W. (1982). Mutagenesis in *Saccharomyces cerevisiae*. *Advan. Genet.* **21**, 173–255.

Leadon, S. A. & Snowden, M. M. (1988). Differential repair of DNA damage in the human metallothionein gene family. *Mol. Cell. Biol.* **8**, 5331–5338.

Lillie, S. H. & Pringle, J. R. (1980). Reserve carbohydrate metabolism in *Saccharomyces cerevisiae*: responses to nutrient limitation. *J. Bacteriol.* **143**, 1384–1394.

Lim, W. A. & Sauer, R. T. (1989). Alternative packing arrangements in the hydrophobic core of λ repressor *Nature (London),* **339**, 31–35.

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

Lörincz, A. T. & Reed, S. I. (1986). Sequence analysis of temperature-sensitive mutations in the *Saccharomyces cerevisiae* gene *CDC28*. *Mol. Cell. Biol.* **6**, 4099–4103.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mellon, I., Spivak, G. & Hanawalt, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian *DHFR* gene. *Cell*, **51**, 241–249.

Melnick, L. & Sherman, F. (1990). Nucleotide sequence of the COR region: a cluster of six genes in the yeast *Saccharomyces cerevisiae*. *Gene*, **87**, 157–166.

Miller, J. H. (1983). Mutational specificity in bacteria. *Annu. Rev. Genet.* **17**, 215–238.

Miyamoto, S., Ohya, Y., Sano, Y., Sakaguchi, S., Iida, H. & Anraku, Y. (1991). A *dbl*-homologous region of the yeast *CLS4/CDC24* gene product is important for Ca^{2+}-modulated bud assembly. *Biochem. Biophys. Res. Commun.* **181**, 604–610.

Mortimer, R. K., Schild, D., Contopoulou, R. & Kans, J. (1989). Genetic map of *Saccharomyces cerevisiae*, edition 10. *Yeast*, **5**, 321–403.

Naumovski, L. & Friedberg, E. C. (1987). The *RAD3* gene of *Saccharomyces cerevisiae*: isolation and characterization of a temperature-sensitive mutant in the essential function and of extragenic suppressors of this mutant. *Mol. Gen. Genet.* **209**, 458–466.

Prakash, L. & Sherman, F. (1973). Mutagenic specificity: reversion of iso-1-cytochrome *c* mutants of yeast. *J. Mol. Biol.* **79**, 65–82.

Pringle, J. R. (1981). The genetic approach to the study of the cell cycle. In *Mitosis/Cytokinesis* (Zimmerman, A. M. & Forer, A., eds), pp. 3–28, Academic Press, New York.

Pringle, J. R. (1987). The gene-number paradox and the complexity of the cell-division cycle and other cellular processes. *Modern Cell Biol.* **5**, 299–306.

Ratzkin, B. & Carbon, J. (1977). Functional expression of cloned yeast DNA in *Escherichia coli*. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 487–491.

Reidhaar-Olson, J. F. & Sauer, R. T. (1988). Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science*, **241**, 53–57.

Rose, M. D., Novick, P., Thomas, J. H., Botstein, D. & Fink, G. R. (1987). A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. *Gene*, **60**, 237–243.

Rothstein, R. J. (1983). One-step gene disruption in yeast. *Methods Enzymol.* **101**, 202–211.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Schatz, P. J., Solomon, F. & Botstein, D. (1988). Isolation and characterization of conditional-lethal mutations in the *TUB1* α-tubulin gene of the yeast *Saccharomyces cerevisiae*. *Genetics*, **120**, 681–695.

Sherman, F., Fink, G. R. & Hicks, J. B. (1986). *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sikorski, R. S. & Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**, 19–27.

Steensma, H. Y., Crowley, J. C. & Kaback, D. B. (1987). Molecular cloning of chromosome I DNA from *Saccharomyces cerevisiae*: isolation and analysis of the *CEN1-ADE1-CDC15* region. *Mol. Cell. Biol.* **7**, 410–419.

Su, J.-Y. & Sclafani, R. A. (1991). Molecular cloning and expression of the human deoxythymidylate kinase gene in yeast. *Nucl. Acids Res.* **19**, 823–827.

Terleth, C., Schenk, P., Poot, R., Brouwer, J. & van de Putte, P. (1990). Differential repair of UV damage in *rad* mutants of *Saccharomyces cerevisiae*: a possible function of G_2 arrest upon UV irradiation. *Mol. Cell. Biol.* **10**, 4678–4684.

Thierry, A., Fairhead, C. & Dujon, B. (1990). The complete sequence of the 8·2 kb segment left of *MAT* on chromosome III reveals five ORFs, including a gene for a yeast ribokinase. *Yeast*, **6**, 521–534.

Wickner, R. B., Koh, T. J., Crowley, J. C., O'Neil, J. & Kaback, D. B. (1987). Molecular cloning of chromosome I DNA from *Saccharomyces cerevisiae*: isolation of the *MAK16* gene and analysis of an adjacent gene essential for growth at low temperatures. *Yeast*, **3**, 51–57.

Wu, T.-H., Clarke, C. H. & Marinus, M. G. (1990). Specificity of *Escherichia coli mutD* and *mutL* mutator strains. *Gene*, **87**, 1–5.

Yoshikawa, A. & Isono, K. (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.

Yoshikawa, A. & Isono, K. (1991). Construction of an ordered clone bank and systematic analysis of the whole transcripts of chromosome VI of *Saccharomyces cerevisiae*. *Nucl. Acids Res.* **19**, 1189–1195.

Edited by S. I. Reed