Construction of a *GAL1*-Regulated Yeast cDNA Expression Library and Its Application to the Identification of Genes Whose Overexpression Causes Lethality in Yeast

Haoping Liu,[1] Janet Krizek and Anthony Bretscher

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853

Manuscript received May 5, 1992
Accepted for publication August 3, 1992

ABSTRACT

We have constructed a galactose-inducible expression library by cloning yeast cDNAs unidirectionally under control of the *GAL1* promoter in a centromeric shuttle vector. Eleven independent libraries were made each with an average size of about 1×10^6 clones, about 50 times larger than the reported mRNA population in a yeast cell. From this library, *LEU2* and *HIS3* cDNAs were recovered at a frequency of about 1 in 10^4 and in 12 out of 13 cases these were expressed in a galactose-dependent manner. Sequence analysis of *leu2* and *his3* complementing cDNAs indicates that they contain all the coding sequence and much of the 5′ untranslated region. To test the utility of the library for the identification of genes whose overexpression confers a specific phenotype, we screened 25,000 yeast transformants for lethality on galactose. Among 15 clones that showed galactose inducible lethality were cDNAs encoding structural proteins, including *ACT1* (actin), *TUB2* (β-tubulin) and *ABP1* (actin-binding protein 1), and genes in signal transduction pathways, including *TPK1* (a cAMP-dependent protein kinase) and *GLC7* (type 1 protein phosphatase). cDNAs overexpressing *NHPB* (nonhistone protein B) and *NSR1* (nuclear sequence recognition protein) were also found to be lethal. Among these, *ACT1* was isolated four times, and *NSR1* three times. The useful features of this library for cDNA cloning in yeast by complementation, and for the identification of genes whose over-expression confers specific phenotypes, are discussed.

THE budding yeast *Saccharomyces cerevisiae* has become one of the most useful organisms in which to undertake genetic and molecular analyses of basic biological processes (BOTSTEIN and FINK 1988). This stems from its ease of manipulation, excellent genetics and its high efficiency of homologous recombination. In particular, mutants with specific defects can be isolated and the relevant genes cloned by complementation of the mutations. To allow for cloning by complementation, a number of genomic libraries have been constructed (ROSE and BROACH 1991). These generally have inserts between 2 and 40 kb and are carried either on a high copy plasmid based on the 2μ circle (NASMYTH and REED 1980; CARLSON and BOTSTEIN 1982) or on a low copy centromeric plasmid (ROSE *et al.* 1987). Analysis of coding sequences from these libraries has often been hampered by their long genomic DNA inserts and the limited availability of useful cloning sites in the vector.

We are interested in the structure and function of the yeast cytoskeleton and have noticed that a number of genes encoding cytoskeletal proteins are lethal when overexpressed. These include the β-tubulin gene (*TUB2*), two genes required for microtubule function (*KAR1* and *BIK1*), and a gene for a nucleoskeletal protein (*NSP1*) (BURKE, GASDASKA and HARTWELL 1989; ROSE and FINK 1987; BERLIN, STYLES and FINK 1990; HURT 1988). Cytoskeletons are dynamic macromolecular assemblies of proteins present in a fairly constant stoichiometric ratio. If one component is synthesized in vast excess, it might thwart the correct assembly and function of the structure. We therefore suspected that the expression level of genes encoding critical components of the cytoskeleton might be important for maintaining their appropriate stoichiometric ratio. This concept predicts that cells might be particularly sensitive to the overexpression of genes encoding these structural components. An earlier example of this line of reasoning (MEEKS-WAGNER and HARTWELL 1986) led these authors to search for genes present in a 2μ genomic library that interfered with mitosis and resulted in a reduced fidelity of chromosome transmission. Using this approach, they isolated two genes (*MIF1* and *MIF2*) whose overexpression enhanced mitotic chromosome loss and transmission (MEEKS-WAGNER *et al.* 1986). Clearly, the over-expression of other genes, particularly regulatory ones such as critical kinases and phosphatases, will also be very deleterious. For these reasons, and to improve the current cloning systems, we set out to construct a highly regulated yeast expression

[1] Present address: Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142.

library and to begin to identify overexpression lethal genes.

Here we describe the construction of a yeast cDNA library whose expression in yeast is regulated by the *GAL1* promoter. The library was cloned into a pBLUESCRIPT-based centromeric plasmid pRS316 which is small, contains many unique restriction sites in the multiple cloning site, and offers versatile functions for DNA manipulations, including easy subcloning, the generation of nested deletions, sequencing, and site directed mutagenesis (SIKORSKI and HIETER 1989). Since the *GAL1* promoter is both strong and tightly regulated, construction of oriented cDNAs next to this promoter should render their expression under the tight control of galactose.

We have characterized the library and begun to analyze genes whose over-expression is lethal. Only a very limited number of genes show over-expression lethality in this system. The library should therefore be useful for gene cloning by complementation. Other potential uses of the library are discussed.

MATERIALS AND METHODS

Strains and media: *Escherichia coli* strains DH5α [*F*$^-$ *endA1 hsdR17*(r_k^-, m_k^+) *supE44 thi-1* λ^- *recA1 gyrA96 relA1* Δ(*argF-laczya*) *U169* ϕ*80 lac* ZΔ*M15*] and DH10B [*F*$^-$ *araD139* Δ(*ara leu*)*7697* Δ*lacX74 galU galK mcrA* Δ(*mrr-hsdRMS-mcrBC*) *rpsL deoR* σ*80dlacZ*Δ*M15 endA1 nupG recA1*] from Bethesda Research Laboratories were used as plasmid hosts. DH10B cells are electrotransformable cells and were used for DNA transformation by electroporation. Yeast strain CUY13 [*MAT***a** *ura3–52*], which originated from the S288C background and was provided by T. HUFFAKER (Cornell University), was used to provide the mRNA for cDNA synthesis. Yeast strains CUY29 [*MAT***a** *ura3–52 leu2–3,112 his3-*Δ*200 lys2 Gal*$^+$] from T. HUFFAKER and YPH226 [*MAT*α *ura3–52 lys2–801 ade2–101 leu2-*Δ*1 his3-*Δ*200 Gal*$^+$] from P. HIETER were used as host strains to test the cDNA library. Media for the growth of yeast were as described by SHERMAN (1991). YEPD contains 1% yeast extract, 2% peptone, 2% glucose; selective media contains yeast nitrogen base with ammonium sulfate and appropriate amino acids containing either 2% glucose or galactose. 5-Fluoroorotic acid (5-FOA) was purchased through the Genetics Society of America and used at a concentration of 0.5 mg/ml (with 5.5 μg/ml uracil) to select against a functional *URA3* gene (BOEKE, LACROUTE and FINK 1984).

Preparation of yeast mRNA and cDNA synthesis: Total RNA from a 4-liter culture of strain CUY14 grown to 5 × 10^7 cells/ml in YEPD was extracted as described by SPRAGUE, JENSEN and HERSKOWITZ (1983). This RNA pellet was stored at −70°. Yeast mRNA was isolated by oligo(dT) cellulose column chromatography (Current Protocols in Molecular Biology) from about 28 mg of total RNA. About 2.8 mg RNA were recovered and precipitated by sodium acetate and ethanol. The oligo(dT) column was regenerated by washing with distilled water, 0.1 M NaOH and water again, and equilibrated with the loading solution. The RNA was rechromatographed on the regenerated oligo(dT) column with about 50% recovery. Then 5 M NaCl was added to the eluted material to bring the salt concentration to 0.5 M so that the RNA could be reapplied to the regenerated oligo(dT) column for the third time. About 50% of the applied RNA was recovered in the elution. After these three cycles through the oligo(dT) column, about 2.5% of RNA was recovered from the total RNA. RNA gels showed that most of tRNA and rRNA had been removed by this procedure and so the material should be highly enriched in mRNA. This mRNA was precipitated with sodium acetate and ethanol and washed with 70% ethanol. It was then redissolved in 1 ml pyrocarbonic acid diethyl ester-treated water and precipitated again to reduce the sodium dodecyl sulfate content in the mRNA, which can interfere with the cDNA synthesis reactions.

Yeast cDNA was synthesized from the mRNA by using the SUPERSCRIPT PLASMID SYSTEM from BRL. This system features directional construction of the library by introducing asymmetry at the ends of cDNA. This is achieved by a primer-adapter, which is a oligo dT primer for first strand synthesis that also contains restriction sites. The primer-adapter used here had the following sequence 5′-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15$^{-3'}$, which contains restriction sites for *Not*I, *Nru*I, *Xba*I and *Spe*I. The 8-base *Not*I recognition sequence next to the 15-bp Ts is extremely rare in most DNAs. All the reactions and procedures were performed as described in the BRL manual. For cDNA synthesis, 3.1 μg mRNA was used, and the yield of the first strand synthesis was 38%. After the second strand cDNA synthesis, *Sal*I adapters were ligated to both ends of the cDNA. The sequence of this adapter is 5′-TCGACCCACGCGTCCG-3′ for the upper strand and 3′-GGGTGCGCAGGCp-5′ for the lower strand. The adapter contains a *MIu*I recognition site and only one of the oligomers in the adapter is phosphorylated, which eliminates self-ligation of the adapters at the *Sal*I overhangs during ligation to the cDNA. The cDNAs were then digested with *Not*I to release the *Not*I sites at the 3′ ends of the cDNAs. The resulting *Sal*I-cDNA-*Not*I fragments (1.23 μg) were size fractionated on a prepacked 1 ml Sephacryl S-500HR column. This is an important step that eliminates a large molar excess of *Sal*I adapters and depletes *Not*I-*Sal*I fragments released from 3′ ends of the cDNAs after the *Not*I digestion. This also reduces the representation of smaller (<500 bp) inserts in the cDNA library. cDNA from fraction numbers 9 (60 ng), 10 (180 ng), 11 (110 ng), and 12 (136 ng) was used to construct the library.

Construction and preparation of vector: The yeast galactose promoter *GAL1* was cloned into the pBLUESCRIPT based yeast CEN plasmid pRS316 (SIKORSKI and HIETER 1989). An *Eco*RI to *Bam*HI DNA fragment carrying the *GAL10* and *GAL1* promoter region was obtained as previously described (LIU and BRETSCHER 1989). The *Bam*HI site is located at 746, 12 bp upstream from the *GAL1* transcription initiation site and the *Eco*RI site is at 1 in the *GAL10* coding sequence (JOHNSTON and DAVIS 1984). The Klenow fragment was used to fill in the ends to which phosphorylated XhoI linkers were added. This DNA fragment was inserted into the *Xho*I site of the pRS316 plasmid and both the *Eco*RI and *Bam*HI sites were retained in this construction. The direction of this insertion is shown in Figure 2. To place cDNAs into the *Sal*I-*Not*I region of the plasmid, plasmid DNA was first digested with *Sal*I to completion. Since there are only 66 bp between the *Sal*I and *Not*I sites, it is difficult to determine whether the second digestion with *Not*I is complete. Three percent of the linearized plasmid DNA was end-labeled with ^{32}P and the DNA was cut with *Not*I. The extent of *Not*I digestion was checked by digesting with *Pvu*II, which has two restriction sites, 186 bp from *Not*I and 1 kb from *Sal*I. Complete digestion by *Not*I will give two ^{32}P-labeled fragments of 1 kb and 66 bp and partial digestion will generate an extra ^{32}P-labeled fragment of 186

TABLE 1

Yeast cDNA library pools

Pool	No. of *E. coli* transformants in the pool
9A	8.2×10^5
9B	6.8×10^5
9C	1.5×10^6
10A	5×10^6
10B	2×10^6
10C	9.4×10^5
10D	1.5×10^6
11A	2.8×10^6
11B	3.2×10^6
12A	4.4×10^6
12B	8.5×10^6

The numbers for each pool correspond to the cDNA fractions from the Sephacryl S-500HR column. The library made from pool 10A is characterized in this paper.

bp. The *SalI-NotI*-cut plasmid was purified away from the 66-bp *SalI-NotI* fragment on a low melt agrose gel (SeaPlaque, FMC Bioproducts) and extracted using an Elutip-D (Schleicher & Schuell, Inc.) column.

Library construction: The cDNA library was constructed as described in the BRL manual except where noted. In a ligation reaction of 20 μl, 10 ng of the cDNA and 50 ng of the *SalI-NotI*-cut plasmid were used. After ligation, the DNA was precipitated and rinsed with 70% ethanol to reduce the salt content. DNA from one ligation was used for one electroporation of 40 μl ElectroMAX DH10B cells (BRL) at 25 μF and 2.5 kV with the pulse controller set to 200 Ω on a Bio-Rad Gene Pulser. The number of transformants from each electroporation is shown in Table 1. Each pool of transformants from 10A and 11A was selected on 100 LB agar plates containing 100 μg/ml ampicillin. Cells were scraped off the plates and resuspended in a sterile solution of 60% LB and 40% glycerol at a concentration of 1×10^{12} cells/ml. Cells were stored at −70°. Other pools of the transformed cells were preserved by adding 1 ml of a sterile solution of 60% LB and 40% glycerol and storing at −70°.

Screening the library for overexpression lethal cDNAs in yeast: About 0.25 μg DNA from the yeast cDNA library was used to transform yeast strains CUY29 or YPH226 by electroporation (BECKER and GUARENTE 1991). The transformation efficiency was about 5×10^5 cells/μg DNA. Cells were plated to generate about 100 colonies per plate and the transformants were selected on SD medium containing 1 M sorbitol and glucose as sole carbon source. They were then replica plated onto selective medium containing either glucose or galactose at 30°. Transformants carrying the putative overexpression lethal genes were tested again on these two media plus a synthetic complete galactose medium containing 5-FOA to select for loss of the cDNA plasmid. This is a critical control as certain strains of yeast became ρ^- and, for unclear reasons, grow extremely slowly or not at all on galactose. To confirm further that the lethality was conferred by the plasmid, the plasmid DNA was recovered from the yeast transformants (STRATHERN and HIGGINS 1991) and transformed into *E. coli* by electroporation. The plasmid was then retransformed into the yeast strain to confirm the galactose dependent lethality. The cDNA in the clones conferring galactose-dependent lethality were partially sequenced by the dideoxynucleotide method (SANGER, NICKLEN and COULSON 1977). The 3′ end of the cDNA was sequenced from the M13 primer (−20) (see Figure 2) and the 5′ end of the gene was sequenced from a primer in the GAL1 promoter region, 690–5′TGCATAACCACTTTA-ACT3′-707, synthesized by the Protein and Nucleic Acid Facility at Cornell University. Sequences were searched by computer for homologies by FASTA databases.

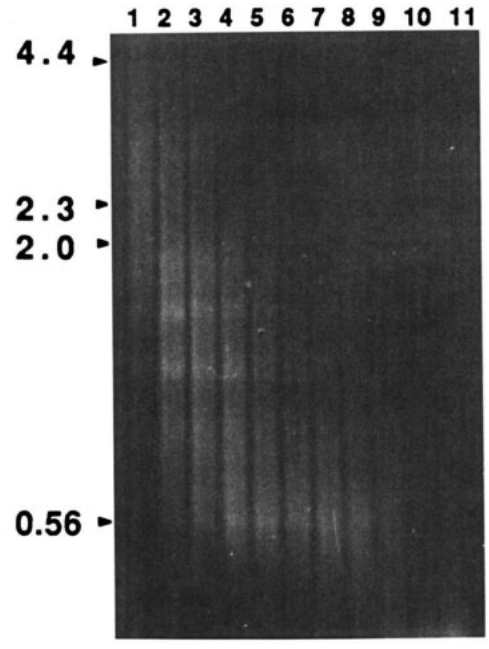

FIGURE 1.—Agarose gel of fractions containing cDNAs from the Sephacryl 500HR column. Lane 1 contains column fractions 9 and 10, lanes 2–11 contain column fractions 11–20, respectively. Fractions 9–12 (lanes 1–3) were used to construct the libraries. Molecular mass markers of λ-*Hin*dIII are shown at left in kilobases.

RESULTS

Construction of the yeast cDNA library: Total RNA was extracted from yeast cells (strain CUY13: *MATa ura3–52*) grown in rich media (YEPD) to reduce the representation of mRNAs for enzymes in biosynthetic pathways. Polyadenylated mRNA was purified from the total RNA by its selective binding to oligo(dT) cellulose. cDNA was synthesized from the mRNA and fractionated on a gel filtration column; Figure 1 shows the size range of the cDNA fractions used for construction of the library. Asymmetry was introduced into the cDNA during its synthesis by constructing a SalI adapter at the end corresponding to the 5′ end of the mRNA and a NotI site at the 3′ end. This enabled us to construct an oriented library by ligating the cDNAs into a yeast CEN plasmid (pRS316-GAL1) that had been digested with *SalI* and *NotI*. This directional ligation of the cDNA into the vector was essential to render expression of the cDNA under the control of the *GAL1* promoter. The map of the library construct is shown in Figure 2. Eleven different cDNA pools were ligated into the plasmid and separately transformed into *E. coli* and stored as independent pools (Table 1). Each pool contained about 10^6 independent transformants. Pool 10A was used to characterize the library further.

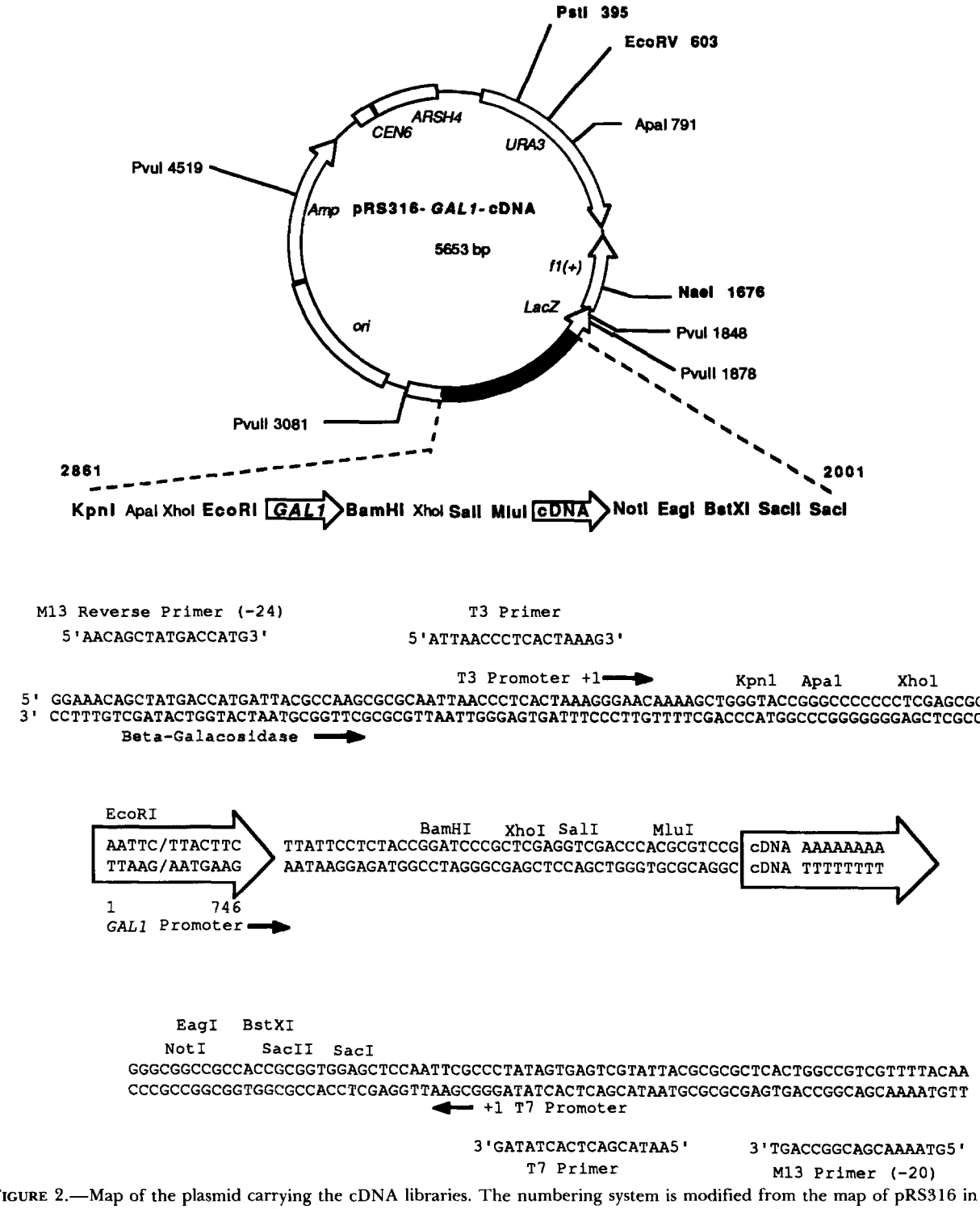

FIGURE 2.—Map of the plasmid carrying the cDNA libraries. The numbering system is modified from the map of pRS316 in SIKORSKI and HIETER (1989). The size of the plasmid does not include the cDNA insert. Enzymes that cleave the vector once or twice are shown; unique restriction sites are shown in **bold** letters. In the lower part of the figure is the DNA sequence surrounding the *GAL1* promoter and the cDNA. Also shown are primers and promoters useful for DNA sequencing and RNA transcription, respectively. This sequence is a combination of the original plasmid sequence of pBLUESCRIPT II KS$^+$ (Stratagene) and direct sequence analysis of of cDNA clones from a primer in the *GAL1* promoter region and the M13 primer (−20).

The cloning strategy employing two different restriction sites gave both an oriented insert and high efficiency of cDNA cloning into the vector. The fraction of plasmids containing insert cDNA was estimated by analyzing plasmid DNA from 57 random *E. coli* transformants. SalI and SacI double digestion of these plasmid DNAs excised cDNA inserts from 55 out of the 57 clones tested, indicating that about 96% of the plasmids contain cDNA inserts. DNA sequencing using the M13 primer from the NotI end into the cDNA (Figure 2), which corresponds to the 3′ end of the mRNA, was done for 8 of these random clones.

TABLE 2

5′ ends of *LEU2* and *HIS3* cDNAs

Gene name	5′ end of cDNA	Major transcription initiation site[a]
HIS3	−2, −4, −7, −12, −13	−12, −23[b]
LEU2	−3, −8, −8, −15, −18, −70	−16[c]

[a] The position of translational initiation is defined as +1.
[b] From STRUHL (1985).
[c] This is a major site, but there appears to be several alternate transcription initiation sites (ANDREADIS *et al.* 1984).

All of the inserts had a minimal stretch of 15 dA residues adjacent to the NotI site, indicating a correct orientation of insertion.

Frequency of specific cDNAs in the library: HEREFORD and ROSBACH (1977) have estimated that *S. cerevisiae* contains 3000–4000 different mRNA sequences that can be classified into low (200 copies/cell), medium (20 copies/cell), and high (1–2 copies/cell) complexity classes. These are estimated to contain approximately 20, 400 and 2400 sequences, respectively. This estimate predicts that the frequency of a single copy mRNA existing in the cDNA library is about 1 in 16,000 clones. To assess the actual frequency of particular cDNAs in the library, the frequency of recovering clones that could complement *his3* and *leu2* mutations was examined. The cDNA library was transformed into a yeast strain (CUY29:*ura5–52 leu2–3,112 his3-Δ200 lys2 GaI*$^+$) selecting for uracil prototrophy and the transformants were replica plated onto selective media containing either glucose or galactose. Of 60,000 transformants, *HIS3* cDNAs were recovered five times and *LEU2* cDNAs eight times. All but one of the clones complemented the amino acid auxotrophy only on medium containing galactose, indicating that 12 out of 13 are under tight control of the *GAL1* promoter. The single clone that did not show galactose-dependent complementation was found to have a *LEU2* cDNA in the inverse orientation. DNA sequencing from the SaII site into the 5′ of the *his3* and *leu2* complementing clones indicated that they were all independent constructs and contained the entire coding regions together with much of the 5′ non-coding sequence (Table 2). It has been estimated that the basal level of *HIS3* gene transcription yields one or two mRNAs per cell under growth conditions where the synthesis of histidine biosynthetic enzymes is repressed (STRUHL 1985; STRUHL and DAVIS 1981). Our finding that *LEU2* cDNAs are recovered at a frequency of about 1 in 7,500 and *HIS3* cDNAs at about 1 in 12,000 is consistent with the predicted figure for the presence of 1–2 copies of their mRNA per cell. The size of each of the eleven independent pools of transformants is therefore about 50-fold larger than the mRNA population. Each pool can be used as an independent library.

Since this is a cDNA library, the majority of the clones would represent the genes of high or moderately abundant mRNAs. Partial DNA sequencing of eight randomly isolated clones identified four full length cDNAs of previously sequenced genes. They are genes for elongation factor 1-α, elongation factor 3, ribosomal protein S7 and ribosomal protein 51B (NAGASHINA *et al.* 1986; QIN *et al.* 1990; ABOVICH and ROSBACH 1984). Yeast containing these plasmids grew well on either glucose or galactose. The other four cDNA sequences did not match sequences in the data bank.

Galactose induced lethal cDNAs: Our original motivation for making this library was to look for genes whose overexpression causes lethality. The cDNA library was transformed into yeast cells and about 25,000 transformants were screened on selective media containing either glucose or galactose as sole carbon source. The transformants that could grow on glucose medium, but not on galactose medium may carry an overexpression lethal cDNA. These plasmids were reisolated and introduced into yeast cells to confirm their galactose-dependent lethality. Fifteen clones retained this lethal phenotype on galactose. The cDNAs in these were sequenced from both ends to identify those derived from known genes. Of the 15, seven were derived from characterized genes, and three from uncharacterized genes. These seven cDNA are derived from transcripts of the following genes: *ACT1* (actin), *TUB2* (β-tubulin), *ABP1* (actin-binding protein), *TPK1* (a cAMP-dependent protein kinase), *NSR1* (nuclear sequence recognition protein), *GLC7* (type 1 protein phosphatase), and *NHP6B* (nonhistone protein B) (GALLWITZ and SURES 1980; NG and ABELSON 1980; NEFF *et al.* 1983; DRUBIN *et al.* 1990; TODA *et al.* 1987; LEE, XUE and MELESE 1991; FENG *et al.* 1991; KOLODRUBETZ and BRUGUM 1990). Among them, *ACT1* cDNAs were isolated four times, and *NSR1* cDNAs were isolated three times (Table 3). One of the open reading frames (ORFs) of unknown function lies between the *BIK1* and *FUS1* genes and its sequence has been reported (TRUEHEART, BOEKE and FINK 1987). Microscopic analysis of galactose induced yeast cells carrying these clones did not reveal any obvious cell cycle-specific arrests.

Quality of the cDNA inserts: The extent of cDNA synthesis for this library was assessed by examining the sequence corresponding to the 5′ ends of the mRNA of the *LEU2* and *HIS3* genes (Table 2), and of the genes in Table 3. *ACT1* mRNA starts at −141 (GALLWITZ, PERRIN and SEIDEL 1981), and the cloned *ACT1* cDNAs start at −97, −98, −99 or −69. NHP6B mRNA starts at −33 and its cDNA at −27. So far we have not encountered any clones that contain partial coding sequences. However, with the exception of the eight randomly selected clones described above, these clones were screened for either their biological func-

TABLE 3

Genes that confer lethality after induction by galactose

Gene[a]	Gene product	Growth on galactose[b]	cDNA starting position[c]
ACT1[d]	Actin	−	−97, −98, −99, −69
TUB2	β-Tubulin	−	−35
ABP1	Actin binding protein I	−	−23
TPK1	cAMP-dependent protein kinase	−	−86
NSR1[e]	Nuclear sequence recognition protein	+/−	−30, −40, −25
GCL7(DIS2S1)	Type 1 protein phosphatase	+/−	−59
NHP6B	Non-histone protein B	+/−	−27
ORF1		−	−23
ORF2		−	
ORF3		−	

[a] DNA sequences of both ends of the cDNAs were sent to FASTA databases to search for homologies. The known genes are listed under their designated names. The other genes are named ORF for open reading frame. ORF1 lies between the *BIK1* and *FUS1* genes and its sequence was published in TRUEHEART, BOEKE and FINK (1987).

[b] No growth after 5 days at 30° on selective media containing galactose is indicated by −, and slight growth by +/−.

[c] The position of translational initiation is defined as +1.

[d] *ACT1* was recovered four times.

[e] *NSR1* was found three times.

tion or galactose inducible lethality and may therefore only represent a fraction of the library. A more accurate assessment is to clone a cDNA from the library by colony hybridization. A single *ACT1* cDNA was independently isolated by hybridization among 2000 *E. coli* colonies screened with labelled *ACT1* DNA. DNA sequence analysis at the 5′ end of the insert showed that the cDNA starts at −105. This *ACT1* cDNA also showed galactose inducible lethality when transformed into yeast cells. This clone, together with the eight randomly selected clones described above, suggest that most of the inserts carry well synthesized cDNAs.

DISCUSSION

We have constructed an oriented yeast cDNA library in a pBLUESCRIPT based yeast centromeric plasmid (Figure 2). This library has a high insert rate (>96%) and appears to contain a high proportion of full length cDNAs. Furthermore, expression of the cDNAs in yeast is under the tight control of the yeast *GAL1* promoter as a consequence of the directional cloning strategy. The number of original transformants in each of 11 pools is about 50 fold in excess of the reported mRNA population in yeast. The frequency of *HIS3* and *LEU2* cDNAs is also in good agreement with the reported mRNA population of about 1 in 16,000. In order to ensure that the entire library is covered, it is probably necessary to screen at least 50,000 transformants, which is easily accomplished by using electroporation to introduce the DNA into yeast.

We see a number of useful features of this library over the currently available cDNA and genomic libraries (MCKNIGHT and MCCONAUGHY 1983; ROSE and BROACH 1991). First, each plasmid contains only the transcribed region of the gene, so there is no need to locate the gene as is the case of cloning from genomic libraries. The fact that the expression of the cDNA is under the control of the *GAL1* promoter may be very useful when cloning by complementation. For example, it will help to rapidly differentiate between real and false clones, which may be especially helpful in the case of complementation of leaky conditional mutations, based on their galactose inducibility. Second, the length of the gene transcript can be estimated directly from the length of the cDNA, which can be excised from the plasmid at the rare restriction sites that flank the cDNA. Third, the sequences of both ends of the cDNA can be readily obtained by using primers in the *GAL1* region and the *lacZ* gene. This sequence information will be sufficient to determine whether the relevant gene has been identified and sequenced previously, which is especially valuable in yeast, where a large percentage of its genes have been sequenced. Fourth, the pBLUESCRIPT polylinker surrounding the cDNA allows unidirectional deletions to be made in the same plasmid. This should be especially useful for the mapping of functional domains as well as in the generation and expression of truncated products. Expression of truncated structural proteins frequently show a dominant negative phenotype, interfering with processes in which the wild-type gene product normally functions (see, for example, HERSKOWITZ 1987; SCHULTHEISS *et al.* 1991). Fifth, the versatile pBLUESCRIPT based vector permits the isolation of single stranded DNA for site specific mutagenesis or sequencing. Its other features include the ability to generate RNA transcripts and to replicate to high copy number in bacteria. All these properties will make characterization of a cDNA cloned from this library relatively fast and easy.

The library was constructed from mRNA expressed in *MAT***a** cells and is therefore limited to genes transcribed in these cells. Genes that are transcribed specifically in diploids or haploid *MATα* cells, or induced under certain growth conditions will therefore not be represented. In addition, mRNAs that have short poly(A) tails may not have been recovered efficiently on the poly-dT affinity column or they might be under represented during the reverse transcription reaction because of the overhang of extra dTs from the primer containing 15 dT. All the cDNA sequences in this library that we have sequenced (>20) end with a poly(A) tail of at least 15 base pairs long. It should be noted that extra dAs might have been added during synthesis of the second strand using the 15 dT of the primer as template.

The library should prove useful for cloning cDNAs by complementation since we have identified only a limited number of over-expression lethal genes in yeast. Moreover, we have found that the introduction of an *ACT1* cDNA clone from the library into an *act1–1* conditional mutant strain can suppress the conditional growth defect in a galactose dependent manner (data not shown). Thus, lethality of an overexpressed cDNA in a wild-type strain does not necessarily imply lethality in the corresponding mutant strain.

Suppression of a conditional mutation in one gene by the overexpression of a different gene has been shown, in many cases, to identify genes whose products are involved in the same functional pathway. For example, cyclin genes were cloned from a 2μ genomic library by their suppression of a *cdc28* mutation (REED *et al.* 1989); and *TPK1* was identified by its suppression of a *cdc25* mutation (TODA *et al.* 1987). This type of analysis should be facilitated using the *GAL1*-cDNA library as it provides two additional features. First, since suppression is galactose-dependent it will facilitate differentiation between suppression from the expressed cDNA and suppression by a chromosomal mutation. Second, once a suppressor cDNA has been identified, it can be analyzed much faster and more easily in this system. In addition, since the level of galactose-induced over-expression of a cDNA on a centromeric plasmid is different from the level of gene overexpression from its own promoter on a multicopy plasmid, the two types of libraries could find complementary uses.

We have used the library to look for genes whose overexpression causes lethality. Of 25,000 clones screened, seven known yeast genes and three unknown genes were found to be lethal following induction on galactose. *ACT1* and *TUB2* genes encode key components of microfilaments and microtubles in yeast. It was previously reported that over-expression of the *TUB2* gene, encoding β-tubulin, caused lethality (BURKE, GASDASKA and HARTWELL 1989), and overexpression of *ACT1*, encoding actin, is lethal (cited in ROSE *et al.* 1987). Isolation of these two classes of cDNAs in our overexpression lethal screen supports the validity of this approach. A five fold overproduction of actin-binding protein I (encoded by *APB1*) on a 2μ plasmid has been reported to render cells temperature sensitive with abnormal actin structures (DRUBIN, MILLER and BOTSTEIN 1988). Here, we have found that expression of this actin-binding protein from the *GAL1* promoter causes cell lethality, possibly because the protein is produced at a higher level in this construct. From these examples and the fact that we have recovered most of the overexpression lethal cDNAs once, it is clear that we have not yet saturated the number of overexpression lethal genes. The fact that we have not yet recovered some genes known to confer over-expression lethality in other systems might be a result of the relative strengths of the promoters used, the plasmid copy number or strain background. However, given the significant fraction of the lethal clones that turned out to encode cDNAs for cytoskeletal proteins, these results support the concept that the expression level of certain cytoskeletal proteins is especially critical for the normal growth of cells.

Other classes of cDNAs were also found to be lethal when overexpressed. The *NSR1* gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs, as well as an acidic N terminus containing a series of serine clusters, and a basic C terminus containing arg-gly repeats (LEE *et al.* 1991). In wild type cells this protein is restricted to a portion of the nucleus. LEE, XUL and MELESE (1991) constructed a *GAL1-NRS1* plasmid that had over 170 bp between the translation initiation codon and the GAL1 promoter, but no overproduction of the protein was observed on galactose. Here, we have cloned the *NSR1* cDNA three times in our screen, suggesting that the *NSR1* mRNA must be quite abundant in the cell. It will be interesting to examine the effect of overproduction of the *NSR1* protein on the localization of nuclear proteins like histone H2B, which contains the nuclear localization sequence recognized by the *NSR1* gene product. *NHP6B* encodes one of the two highly homologous nonhistone proteins in yeast (KOLODRUBETZ and BRUGUM 1990). Both these non-histone proteins are homologous to the middle segment of the chromatin-associated high mobility group protein 1 from calf and are speculated to have a role in transcription or DNA replication.

Another group of the clones identified in our overexpression lethal screen encode a protein kinase (*TPK1*) and a protein phosphatase (*GLC7*). *TPK1* encodes one of the three catalytic subunits of the cAMP-dependent protein kinase in yeast; however, the TPK1

gene on a 2μ plasmid does not cause cell lethality (TODA *et al.* 1987). *GLC7* encodes a type 1 protein phosphatase required for glycogen accumulation in yeast (FENG *et al.* 1991). *GLC7* is identical to the *DIS2S2* which was cloned by its homology to the *Schizosaccharomyces pombe DIS2* (defective in sister chromatid disjoining) gene, believed to be involved in mitosis (OHKURA *et al.* 1991). FENG *et al.* (1991) reported that *GLC7* mRNA increased 4-fold at the end of exponential growth in wild type cells with the implication that activation of glycogen synthesis is mediated by increased expression of protein phosphatase 1 as cells reach stationary phase. If this is true, we would expect to see an increased amount of glycogen synthesis in cells overexpressing *GLC7*; however, in preliminary experiments no unusual accumulation of glycogen was found in these cells after induction by galactose (data not shown).

Among the three unknown overexpression lethal genes, ORF1 is the gene between *BIK1* and *FUS1*. Its deduced protein sequence consists largely of glutamines and aspargines, as well as a stretch of 20 alternating glutamine and glycine residues (TRUEHEART, BOEKE and FINK 1987). Remarkably, deletion of this gene has no obvious phenotype, yet we find that overexpression confers lethality. This finding underscores the rationale for using divergent genetic approaches to identify genes whose products are involved in specific processes.

We believe that our results point to two important uses of this library. The first is that it will make cloning by complementation and suppression analysis by overexpression easier. Second, and more importantly, it provides a genetic approach of identifying genes whose overexpression interferes with specific biological processes. In our studies we have examined clones that confer lethality in a galactose-dependent manner. It may be equally useful to select clones that overcome normally growth-inhibiting conditions, such as the G_1 arrest imposed on *MAT***a** cells by α-factor. In this example, overexpression of cDNAs that encode components of the signal transduction pathway might interfere with correct transmission of the signal that arrests the cells in G_1 and thereby allow growth in the presence of α-factor. This type of approach and our lethal screen make use of the most extreme phenotypes, namely induced life or induced death. However, more subtle screens or selections might also uncover cDNAs that encode important proteins. For example, this was the basis for the isolation of the *MIF1* and *MIF2* (MEEKS-WAGNER *et al.* 1986) genes discussed above. The regulated cDNA expression library should facilitate this type of approach and together with the classic conditional lethal mutation analysis allow for a more complete understanding of genes whose products play roles in specific processes.

We are grateful to TOM FOX, BIK TYE and to members of this laboratory for discussions. Special thanks go to ELIZABETH CHEN for computer expertise and for isolating an *ACT1* cDNA by colony hybridization and showing that it confers overexpression lethality, to DANIEL SCHOTT who used this cDNA to show that it can rescue the *act1-1* conditional mutation at its restrictive temperature, and to RUBY YE who looked for glycogen accumulation in cells overexpression *GLC7*. This work was supported by grant GM39066 from the National Institutes of Health.

LITERATURE CITED

ABOVICH, N., and M. ROSBACH, 1984 Two genes for ribosomal protein 51 of *Saccharomyces cerevisiae* complement and contribute to the ribosomes. Mol. Cell Biol. **4:** 1871–1879.

ANDEADIS, A., Y.-P. HSU, M. HERMODSON, G. KOHLHAW and P. SCHIMMEL, 1984 Yeast *LEU2*. Repression of mRNA levels by leucine and primary structure of the gene product. J. Biol. Chem. **259:** 8059–8062.

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

BERLIN, V., C. STYLES and G. R. FINK, 1990 *BIK1*, a protein required for microtubule function during mating and mitosis in *Saccharomyces cerevisiae*, colocalizes with tubulin. J. Cell Biol. **111:** 2573–2586.

BOEKE, J. D., F. LACROUTE and G. R. FINK, 1984 A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol. Gen. Genet. **197:** 345–347.

BOTSTEIN, D., and G. R. FINK, 1988 Yeast as an experimental organism for modern biology. Science **240:** 1439–1443.

BURKE, D., P. GASDASKA and L. HARTWELL, 1989 Dominant effects of tubulin overexpression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **9:** 1049–1059.

CARLSON, M., and D. BOTSTEIN, 1982 Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. Cell **28:** 145–154.

DRUBIN, D. G., K. G. MILLER and D. BOTSTEIN, 1988 Yeast actin-binding proteins: evidence for a role in morphogenesis. J. Cell Biol. **107:** 2551–2561.

DRUBIN, D. G., J. MULHOLLAND, Z. M. ZHU and D. BOTSTEIN, 1990 Homology of a yeast actin-binding protein to signal transduction proteins and myosin-I. Nature **343:** 388–390.

FENG, Z., S. E. WILSON, Z.-Y. PENG, K. K. SCHLENDER, E. M. REIMANN and R. J. TRUMBLY, 1991 The yeast GLC7 gene required for glycogen accumulation encodes a type 1 protein phosphatase. J. Biol. Chem. **266:** 23796–23801.

GALLWITZ, D., F. PERRIN and R. SEIDEL, 1981 The actin gene in yeast *Saccharomyces cerevisiae:* 5′ and 3′ end mapping, flanking and putative regulatory sequences. Nucleic Acids Res. **9:** 6339–6350.

GALLWITZ, D., and I. SURES, 1980 Structure of a split gene: complete nucleotide sequence of the actin gene in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **77:** 2546–2550.

HEREFORD, L. M., and M. ROSBACH, 1977 Number and distribution of polyadenylated RNA sequences in yeast. Cell **10:** 453–462.

HERSKOWITZ, I., 1987 Functional inactivation of genes by dominant negative mutations. Nature **329:** 219–222.

HURT, E. C., 1988 A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of *Saccharomyces cerevisiae*. EMBO J. **7:** 4323–4334.

JOHNSTON, M., and R. W. DAVIS, 1984 Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. Mol. Cel. Biol. **4:** 1440–1448.

KOLODRUBETZ, D., and A. BURGUM, 1990 Duplicated NHP6 genes of *Saccharomyces cerevisiae* encode proteins homologous

to bovine high mobility group protein. 1. J. Biol. Chem. **265:** 3234–3239.

LEE, W.-C., Z. XUE and T. MELESE, 1991 The NSR1 gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs. J. Cell. Biol. **113:** 1–12.

LIU, H., and A. BRETSCHER, 1989 Disruption of the single tropomyosin gene in yeast results in the disappearance of actin cables from the cytoskeleton. Cell **57:** 233–242.

MCKNIGHT, G. L., and B. L. MCCONAUGHY, 1983 Selection of functional cDNAs by complementation in yeast. Proc. Natl. Acad. Sci. USA **80:** 4412–4416.

MEEKS-WAGNER, D., and L. H. HARTWELL, 1986 Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell **44:** 43–52.

MEEKS-WAGNER, D., J. S. WOOD, B. GARVIK and L. H. HARTWELL, 1986 Isolation of two genes that affect mitotic chromosome transmission in *S. cerevisiae*. Cell **44:** 53–63.

NAGASHINA, K., M. KASAI, S. NAGATA and Y. KAZIRO, 1986 Structure of the two genes coding for polypeptide chain elongation factor 1-alpha (EF-1-alpha) from *Saccharomyces cerevisiae*. Gene **45:** 265–273.

NASMYTH, K., and S. I. REED, 1980 The isolation of genes by complementation in yeast: the molecular cloning of a cell cycle gene. Proc. Natl. Acad. Sci. USA **77:** 2119–2123.

NEFF, N. F., J. H. THOMAS, P. GRISAFI and D. BOTSTEIN, 1983 Isolation of the beta-tubulin gene from yeast and demonstration of its essential role in vivo. Cell **33:** 211–219.

NG, R., and J. ABELSON, 1980 Isolation and sequence of the gene for actin in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **77:** 3912–3916.

OHKURA, H., N. KINOSHITA, S. MIYATANI, T. TODA and M. YANAGIDA, 1991 The fission yeast *dis2*$^+$ gene required for chromosome disjoining encodes one of two putative protein phosphatases. Cell **57:** 997–1007.

QIN, S., A. XIE, M. C. BONATO and C. S. MCLAUGHLIN, 1990 Sequence analysis of the translational elongation factor 3 from *Saccharomyces cerevisiae*. J. Biol. Chem. **265:** 1903–1912.

REED, S. I., J. A. HADWIGER, H. E. RICHARDSON and C. WITTENBERG, 1989 Analysis of the *Cdc28* protein kinase complex by dosage suppression. J. Cell Sci. Suppl. **12:** 29–37.

ROSE, M. D., and J. R. BROACH, 1991 Cloning genes by complementation in yeast. Methods Enzymol. **194:** 195–230.

ROSE, M. D., and G. R. FINK, 1987 *KAR1*, a gene required for function of both intranuclear and extranuclear microtubules in yeast. Cell **48:** 1047–1060.

ROSE, M. D., P. NOVICK, J. H. THOMAS and G. R. FINK, 1987 A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. Gene **60:** 237–243.

SANGER, F., NICKLEN, S. and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SCHULTHEISS, T., Z. LIN, H. ISHIKAWA, I. U. ZAMIR, C. STOECKERT and H. HOLTZER, 1991 Desmin/vimentin intermediate filaments are dispensible for many aspects of myogenesis. J. Cell Biol. **114:** 953–966.

SHERMAN, F., 1991 Getting started with yeast. Methods Enzymol. **194:** 3–21.

SIKORSKI, R. S., and P. HIETER, 1989 A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics **122:** 19–27.

SPRAGUE, G. F., JR., R. JENSEN and I. HERSKOWITZ, 1983 Control of yeast cell type by the mating type locus: positive regulation of the α-specific *STE3* gene by the *MATα1* product. Cell **32:** 409–415.

STRATHERN, J. N., and D. R. HIGGINS, 1991 Recovery of plasmids from yeast into *Escherichia coli:* shuttle vectors. Methods Enzymol. **194:** 319–329.

STRUHL, K., 1985 Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. Nucleic Acids Res. **13:** 8587–8601.

STRUHL, K., and R. DAVIS, 1981 Transcription of the *his3* gene region in *Saccharomyces cerevisiae*. J. Mol. Biol. **152:** 535–552.

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.

TODA, T., S. CAMERON, P. SASS, M. ZOLLER and M. WIGLER, 1987 Three different genes in *S. cerevisiae* encode the catalytic subunits of the cAMP-dependent protein kinase. Cell **50:** 27–287.

Communicating editor: A. G. HINNEBUSCH