Dominant genetics using a yeast genomic library under the control of a strong inducible promoter

(SYES/expression library/truncated proteins/mating response pathway/STEII)

SANDRA W. RAMER, STEPHEN J. ELLEDGE*, AND RONALD W. DAVIS

Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305

Contributed by Ronald W. Davis, September 1, 1992

ABSTRACT In Saccharomyces cerevisiae, numerous genes have been identified by selection from high-copy-number libraries based on "multicopy suppression" or other phenotypic consequences of overexpression. Although fruitful, this approach suffers from two major drawbacks. First, high copy number alone may not permit high-level expression of tightly regulated genes. Conversely, other genes expressed in proportion to dosage cannot be identified if their products are toxic at elevated levels. This work reports construction of a genomic DNA expression library for S. cerevisiae that circumvents both limitations by fusing randomly sheared genomic DNA to the strong, inducible yeast GALI promoter, which can be regulated by carbon source. The library obtained contains 5 x 10^7 independent recombinants, representing a breakpoint at every base in the yeast genome. This library was used to examine aberrant gene expression in S. cerevisiae. A screen for dominant activators of yeast mating response identified eight genes that activate the pathway in the absence of exogenous mating pheromone, including one previously unidentified gene. One activator was a truncated STEII gene lacking ~1000 base pairs of amino-terminal coding sequence. In two different clones, the same GALI promoter-proximal ATG is in-frame with the coding sequence of STEII, suggesting that internal initiation of translation there results in production of a biologically active, truncated STEII protein. Thus this library allows isolation based on dominant phenotypes of genes that might have been difficult or impossible to isolate from high-copy-number libraries.

A common genetic approach employed to analyze the cellular function of a protein is an examination of the phenotypes associated with perturbing its expression level. The effects of a protein's absence can be studied readily in Saccharomyces cerevisiae by using homologous recombination to delete genes (1). In addition, many proteins show novel phenotypes when overproduced or inappropriately produced. Increased gene dosage is sometimes sufficient to produce elevated levels of a given protein (2). High-copy-number libraries are used frequently in yeast to isolate genes based on phenotypes due to increased expression levels. However, these libraries suffer from two major drawbacks due to their dependence on simple gene dosage to elevate protein production. First, high levels of protein production cannot always be achieved by high copy number alone (3). Classes of genes such as those expressed at specific points in the cell cycle or during particular growth or differentiation stages, for example, may be tightly regulated and thus not subject to overexpression by increased gene dosage. Thus many genes that might show interesting phenotypes if overexpressed cannot be isolated in this manner. Second, because these libraries lack a mechanism for controlling protein production, they are deficient in genes whose products are toxic at high concentration and in genes tightly linked to these toxic genes.

To circumvent these limitations, we have constructed a yeast genomic DNA expression library consisting of random genomic fragments cloned downstream of the regulatable GALI promoter (PGALI). Thus a portion of the library contains coding sequences that have been separated from their endogenous promoter regions and joined instead to PGALI. PGALI is a very strong yeast promoter that is tightly repressed in cells utilizing glucose as a carbon source and is induced over 1000-fold when the carbon source is changed to galactose (4). Thus clones from the library that contain PGALI fusions proximal to coding sequences can achieve high level expression of these genes in an experimentally regulatable manner. Because genes can be isolated based on the dominant phenotypes displayed when they are overexpressed, this method is termed "gene isolation by conditional overexpression," and the regulatable library is referred to as a "conditional expression library."

The library was tested by a variety of standard criteria, and the ability of clones in the library to complement amino acid auxotrophies in both a galactose-dependent and -independent fashion was confirmed. Two additional screens were undertaken to examine the ability of the features of this library to permit the isolation of genes with interesting phenotypes. The first screen identified a gene whose protein product is deleterious at high concentration.

The second screen was designed to identify genes that show novel phenotypes when overexpressed. Two such genes that have been identified previously in yeast, STE4 and STE12, are involved in pheromone-induced mating response. Overexpressing either gene from PGALI can cause the mating-response pathway in the absence of exogenous mating pheromone (5–8). The conditional expression library was screened for other genes that could activate the pathway when overproduced. The identification of several such genes, taken together with the other uses of the library, suggests that this regulatable genomic DNA expression library will be a valuable alternative to standard high-copy-number libraries in S. cerevisiae.

MATERIALS AND METHODS

Plasmids, Strains, and Media. Plasmids and strains used for library construction have been described elsewhere in detail (9). The features of the pYES-R vector relevant to this work are diagrammed in Fig. 1. Genomic yeast DNA was prepared from strain SNY243 (MATa leu2-04 ade1 ade6 cir6; from M. Snyder, Yale University). Other yeast strains used were YC3

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. $1734 solely to indicate this fact.

Abbreviation: PGALI, GALI promoter.

*Present address: Department of Biochemistry and Institute for Molecular Genetics, Baylor College of Medicine, Houston, TX 77030.

The sequence reported in this paper has been deposited in the GenBank data base (accession no. L02617).
A yeast genomic DNA library was prepared in a vector designed to allow conditional, high-level protein production in yeast resulting from the fusion of coding sequences to a strong, vector-borne promoter. The resulting library was large and contained a high proportion of DNA inserts. The ability of the library to allow the identification of clones that would not be found in a traditional high-copy-number library was examined. Several such clones were isolated, indicating that this conditional expression library can be a useful alternative to such high-copy-number libraries.

**Preparation of a Yeast Genomic DNA Library in pYES-R.** A library of 5 × 10⁷ recombinant phage was constructed in *AYES*-R (9). An aliquot of the amplified library representing 5 × 10⁷ phage was converted to the plasmid form (pYES-R) by using the “automatic subcloning” feature of the *AYES* system (9). A library of this complexity contains enough clones to represent fusion of *P*₉₆ in most every base in the yeast genome (14 million bp). Nine of 10 randomly selected, independent library clones analyzed contained genomic DNA inserts, with an average insert size of 3800 bp.

**Complementation of Amino Acid Auxotrophy.** To verify that plasmids in the library contain genes under the control of *P*₉₆, the library was screened for clones able to complement the histidine auxotrophy of a *his3Δ200* strain, YC3, in a galactose-dependent fashion. Out of 100 000 YC3 transformants, 42 clones were capable of growing on SGC-His medium. Four of these clones could not grow on SDC-His medium, indicating that the *HIS3* gene was under the control of *P*₉₆. Plasmid DNA was rescued from one galactose-dependent HIS⁺ clone. The *P*₉₆ proximal end of the genomic insert DNA from this plasmid was sequenced. The site of the fusion was 59 bp upstream of the translational start codon of the *HIS3* gene (Fig. 2A).

**Isolation of a Gene Conferring Galactose-Dependent Growth Arrest.** The ability of *P*₉₆ to be regulated experimentally by choice of carbon source allows plasmids from the library to be maintained under conditions in which there is no *P*₉₆-driven transcription. This feature should permit the isolation of genes whose products are toxic at high concentration (Toxic on High Expression, or *THE* genes). Such genes could not be represented in a standard high-copy-number library. To identify such *THE* genes, 40,000 transformants in YNN490 were plated at low density (<500 per plate) and replica-plated to glucose plates and to galactose plates. Visual comparison of the two sets of plates revealed several colonies whose growth on galactose plates was much poorer. When streaked for single colonies, all but one of these transformants proved capable of growing on galactose, although much more slowly than the majority of the library.

**RESULTS**

ml of ice-cold Z buffer containing 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 2 mM pepstatin A, 0.6 mM leupeptin, and 0.5 mM benzamidine by using glass beads (13), and β-galactosidase activity was determined by standard techniques (14). For assays requiring galactose induction, logarithmic-phase cells in glucose-containing media were pelleted by centrifugation, washed with sterile water, and then resuspended in galactose-containing media for 17 h before extracts were prepared.

**DNA Sequencing.** Double-stranded plasmid DNA was sequenced by the method of Sanger et al. (15) by using Sequenase 2.0 as recommended by the manufacturer (United States Biochemical). The oligonucleotide 5'-CAACAAAA- AATTGTTAATATACCTC-3', corresponding to a region of the yeast *P*₉₆ was used as a primer to obtain sequence information from genomic insert DNA proximal to *P*₉₆. Sequence analysis and homology searching were performed using the Intelligenetics programs.

**FIG. 1.** Features of the pYES-R vector used for genomic library construction. Details of the vector have been described (9).

(MATα ade6 can1-100 his3-11,15 leu2-3,112 trpl-1 ura3-1; from R. Esposito, University of Chicago), CRY1 (MATα can1-100 ade2-1 his3-11,15 leu2-3,112 trpl-1 ura3-1; from R. Fuller, Stanford University), YNN490 (MATα/a ade2-1/+ fade6 can1-100/his3-11,15/his3-11,15 leu2- 3,112/leu2-3,112 trpl-1/trpl-1 ura3-1/ura3-1), YNN217 (MATα lys2-801 ura3-52 ade2-101 his3Δ200), and YNN19 (MATα FUS1-lacZ::LEU2 ura3 leu2 his3 trpl 1 lys2; ref. 6). Standard media for yeast and Escherichia coli were used as described (10, 11). Uracil-deficient media were used to select for the pYES library plasmids.

**Preparation of a S. cerevisiae Genomic DNA Library.** Yeast genomic DNA from strain SNY243 was purified and prepared as described (12). The DNA was sheared by passage through a 27-gauge needle until the average size was ~6000 base pairs (bp). The sheared DNA was made blunt by using T4 DNA polymerase. Phosphorylated *Xho I* adaptors (5'-CGAGTCACGTCAAGGG-3' and 5'-CCCTGACGTAGC-3') were ligated to the repaired genomic DNA at a 10-fold molar excess over genomic DNA. Fragments of 3600-8400 bp were gel purified, ligated to "T-filled" AYES-R arms, and packaged as described (9). The library was amplified by one passage through *E. coli* (11).

The phage library was converted to a plasmid library suitable for transformation into yeast by using *cre-lox* recombination essentially as described (9). Fifty million plaque-forming units, equal to the number of independent clones in the initial unamplified library, yielded 4 mg of cesium chloride-purified plasmid DNA.

**Complementation of his3⁰ Auxotrophy.** Transformants were selected on SDC-Ura (10) plates and replica-plated to SGC-His (10) plates. Colonies that grew on the SGC-His medium were tested for growth on SDC-His (10) plates.

**Filter Replica Assay for β-Galactosidase Activity.** Nitrocellulose filter replicas were made from plates containing freshly grown yeast colonies, and the yeast cells were permeabilized by immersing the filter in liquid nitrogen for 5–10 sec. Filters were thawed and then placed in a Petri dish colony-side-up on Whatman 3MM paper saturated with Z buffer (60 mM Na₂HPO₄/40 mM NaH₂PO₄/10 mM KCl/1 mM MgSO₄/50 mM 2-mercaptoethanol) containing 5-bromo-4-chloro-3-indolyl-β-d-galactoside (300 μg/ml). Filters were incubated at 30°C for between 30 min and 18 h and then scored for blue colonies. For determining galactose-dependent β-galactosidase activity, filters were incubated colony-side-up on plates containing galactose as the sole carbon source for 12–18 h prior to assay.

**Quantitative β-Galactosidase Assays.** A crude protein extract was prepared from 10⁶ late-logarithmic-phase cells in 1
transforms. The final transformant produced only occasional papillae when streaked onto galactose plates but grew well on glucose medium. The library plasmid was rescued from this clone and was used to transform several other strains (YNN490, YNN217, and NNY19). In each case, the plasmid conferred galactose-dependent growth arrest, indicating that the effect is neither strain nor cell-type specific. Galactose-induced cells of all three strains containing the plasmid were mononucleate, showed no gross morphological abnormalities, and did not arrest with a CDC phenotype (data not shown).

Sequence data from the end of the genomic insert DNA proximal to P\textsubscript{GALI} revealed an open reading frame \textit{(THEI)} with no similarity to other sequences presently in the DNA data bases (Fig. 2B). Sequence data from the other end of the genomic insert DNA, nearest the \textit{lac} promoter, indicated that a portion of the \textit{ADE2} gene was present in the insert fragment, thus mapping \textit{THEI} to the long arm of chromosome XV, within 3000 bp of \textit{ADE2}.

**Screen for Activators of the Mating Response Pathway.**

Activation of the pheromone response pathway in the absence of pheromone is a phenotype exhibited by the \textit{STE4} and \textit{STE12} genes, two of the genes involved in the pathway, when overproduced under \textit{P\textsubscript{GALI}} control (5–8). To test for the presence in the library of other such genes, transformants in NNY19 were screened for the ability to induce transcription of \textit{FUSI}, a gene that is induced over 40-fold soon after cells are exposed to mating pheromone. \textit{FUSI} transcription has been used often as a measure of primary intracellular response to pheromone exposure (16). A chromosomally integrated copy of a \textit{FUSI}–lacZ fusion was used as a reporter for \textit{FUSI} transcriptional activation (6). Thirteen of 150,000 transformants screened conferred plasmid-dependent, galactose-dependent activation of the \textit{FUSI} fusion not requiring exogenously supplied mating pheromone. The \textit{P\textsubscript{GALI}} proximal ends of the genomic inserts in these plasmids were sequenced, and the sequence data were used to identify 12 of these clones. One clone was isolated that did not match sequences available in the DNA data bases. This clone is being analyzed further, and its characterization will be described elsewhere. The remaining clones corresponded to previously identified genes (Table 1).

The positions of the vector-insert fusion junctions for the identified genes are listed in Table 1. Three of the genes, \textit{OBFI}, \textit{PHO2}, and \textit{STE11}, were isolated twice. The site of fusion was unique for each of these clones, indicating that they were independently derived. Three \textit{MATa1} clones were isolated, each with a unique fusion junction ranging from 34 to 84 bases upstream of the initiation codon. \textit{P\textsubscript{GALI}}-gene fusions occurred as close as 18 bases upstream of the initiation codon to as far as 256 bases upstream of the ATG. The average distance from the start codon for all the known genes obtained is 93 bp.

In the case of two of the \textit{MATa1} clones and the \textit{MSNI} isolate, the fusion points each occurred upstream of two ATGs, neither of which was in-frame with the coding sequence. The leader sequences of the remaining genes contained no intervening ATGs.

The two \textit{STE11} isolates were particularly interesting. Neither clone contained the full-length \textit{STE11} gene. The insert fragment in each clone contained only the carboxyl-terminal coding region for \textit{STE11} plus sequences further downstream of the gene. In each case, approximately half of the \textit{STE11} coding region was absent. In one clone, the site of fusion with \textit{P\textsubscript{GALI}} was located 934 bp into the coding sequence of \textit{STE11}, and in the other it was 953 bp into the gene. In both cases, the \textit{P\textsubscript{GALI}}-gene fusion occurred such that the first ATG after the junction was in-frame with the \textit{STE11} gene coding sequence. With leader sequences of 68 and 87 bp, the two clones putatively initiate translation from the same internal, in-frame ATG. One of the truncation alleles has been characterized further in a study described elsewhere (17).

**Quantification of Transcriptional Activation of \textit{FUSI}–lacZ.**

To assess the sensitivity of the visual screen and to determine the relative levels of \textit{FUSI}–lacZ induction elicited by over-expression of each clone, quantitative \textit{β}-galactosidase assays were performed. Table 1 lists the levels of activity produced by cells containing each of the clones after 17 h of growth in galactose-containing media. As little as a 3-fold increase in \textit{β}-galactosidase activity was perceptible in the visual screen.

**DISCUSSION**

The development of an experimentally regulatable yeast genomic DNA expression library was prompted by a consideration of the advantages such a library would have over standard high-copy-number libraries. The ability to experimentally control protein production enables studies of inappropriate gene expression that cannot be undertaken using high-copy-number libraries. Such studies could include, for example, the \textit{obfl} gene.

**Table 1. Results of screen for activators of mating response pathway**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fusion junction point, bp upstream of initiation codon</th>
<th>(β)-Galactosidase, fold induction*</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{OBFI}</td>
<td>105</td>
<td>3 ± 0.2</td>
</tr>
<tr>
<td>\textit{OBFI}</td>
<td>20</td>
<td>7 ± 0.3</td>
</tr>
<tr>
<td>\textit{MSNI}</td>
<td>151</td>
<td>4 ± 2</td>
</tr>
<tr>
<td>\textit{MATa1}</td>
<td>68</td>
<td>13 ± 0.8</td>
</tr>
<tr>
<td>\textit{MATa1}</td>
<td>36</td>
<td>27 ± 1</td>
</tr>
<tr>
<td>\textit{MATa1}</td>
<td>84</td>
<td>49 ± 3</td>
</tr>
<tr>
<td>\textit{PHO2}</td>
<td>135</td>
<td>36 ± 3</td>
</tr>
<tr>
<td>\textit{PHO2}</td>
<td>18</td>
<td>52 ± 3</td>
</tr>
<tr>
<td>\textit{MCNI}</td>
<td>256</td>
<td>44 ± 5</td>
</tr>
<tr>
<td>\textit{STE4}</td>
<td>84</td>
<td>1600 ± 100</td>
</tr>
<tr>
<td>\textit{STE11}</td>
<td>68†</td>
<td>2800 ± 300</td>
</tr>
<tr>
<td>\textit{STE11}</td>
<td>87†</td>
<td>3000 ± 200</td>
</tr>
</tbody>
</table>

*Fold induction relative to the parent strain without a plasmid, grown on galactose. Values are the average (± SD) of two trials.
†Position of fusion junction point upstream of the internal ATG (corresponding to amino acid 382 of the full-length protein) putatively used as start codon for truncation protein.
example, investigating the effects of constitutive expression of genes that are normally temporally regulated with the cell cycle, expression in all cell types of genes that are normally expressed only in certain cell types (e.g., haploid- or diploid-specific genes), or expression of genes in inappropriate growth states of the cell (e.g., mating, meiosis, or sporulation-specific genes during vegetative growth). The conditional expression library should be able to increase the protein level of virtually any yeast gene above its wild-type level, because P\textsubscript{GAL1} is one of the strongest yeast promoters identified (4). Because cells containing the library can be maintained under conditions where transcription from P\textsubscript{GAL1} is repressed, the conditional expression library should also permit identification of genes whose protein products are toxic at high levels. Such clones should be absent from a high-copy-number library because their unregulated overexpression should be lethal. Finally, because the conditional overexpression described here is plasmid-based, the library can be used to study dominant effects of overexpression without the attendant genetic difficulties often associated with studying dominant phenotypes. Thus, the potential advantages of a conditional expression library over a traditional high-copy-number library are manifold.

To achieve regulatable, high-level expression of a given gene under the strict control of P\textsubscript{GAL1}, two criteria must be met. First, P\textsubscript{GAL1} should be fused to a region between the coding sequence and its endogenous promoter and upstream regulatory regions. Second, there should not be fortuitous ATG start codons between P\textsubscript{GAL1} and the bona fide start of the gene, because translation in yeast usually begins at the first ATG in a given message (18). These theoretical considerations and an examination of the upstream regions of several sequenced yeast genes suggested that a P\textsubscript{GAL1}-gene fusion occurring up to 100 bp upstream of most coding sequences should result in regulatable control by P\textsubscript{GAL1}. For a haploid genome of ≈14 million bp, a minimal library size of at least 140,000 clones, representing a breakpoint every 100 bases, would be necessary to achieve a functional P\textsubscript{GAL1}-gene fusion for each gene in the yeast genome.

The library obtained in this study is complex enough to contain clones representing a breakpoint at every base in the genome, assuming sequence-independent shearing. Thus the library should contain every possible P\textsubscript{GAL1}-gene fusion. However, P\textsubscript{GAL1}-gene fusions that meet the criteria described above will result in P\textsubscript{GAL1}-regulated protein production. The majority of the clones in the library contain genes that retain their endogenous regulatory sequences, and thus the library functions independently of carbon source as a standard single-copy library as well. Additionally, because the library was made in the pYES vector and the genomic DNA was inserted nondirectionally, the library should also contain genes that can be expressed in E. coli by using the lac promoter. In its phage form, it should be suitable for screening with antibodies as well.

The isolation of clones capable of complementing amino acid auxotrophies in both a galactose-independent and galactose-dependent fashion is a mainstay of yeast genetics. The library contains both endogenously regulated genes and coding sequences under P\textsubscript{GAL1} control. Endogenously regulated HIS3 genes were isolated at approximately the frequency expected for a standard genomic DNA library (19). In addition, four clones were isolated that could complement a his\textsubscript{3} mutation in a strict galactose-dependent manner. To achieve strict galactose-dependent transcription, the P\textsubscript{GAL1} fusion point must usually occur close enough to the coding sequence to start the endogenous regulatory signals. In the galactose-dependent HIS3 clone that was sequenced, the fusion junction occurred immediately downstream of the HIS3 TATA boxes (20), thus putatively eliminating endogenous HIS3 regulation. Interestingly, however, it occurred upstream of two out-of-frame ATGs, which might have been predicted to interfere with P\textsubscript{GAL1}-driven HIS3 expression. Three clones identified in the mating-response screen contained intervening out-of-frame ATGs as well. Thus in at least some cases, the presence of intervening ATGs between P\textsubscript{GAL1} and the coding sequence initiation codon does not preclude expression of gene products under galactose control.

The set of genes whose RNA or protein products are deleterious at high dosage represents an important class of genes that cannot readily be isolated by using existing techniques. Although there have been reports of specific genes that were lethal when cloned into an overexpression system (21), these genes cannot generally be found by using a high-copy-number library. Thus it has not been feasible to identify more than isolated examples of such genes. Gene isolation by conditional overexpression using this library should permit the identification of a whole class of these genes, and, indeed, screening a small fraction of the library for genes that are toxic on overexpression readily identified one such gene, THE1.

A conditional overexpression library also makes possible the study of dominant overexpression phenotypes. Two genes in the yeast pheromone response pathway, STE4 and STE12, cause dominant, pheromone-independent activation of FUS1 transcription when overproduced from P\textsubscript{GAL1} (5, 6, 8). In the current study, 150,000 library transformants were screened for dominant activation of the mating response pathway. Analysis of the positive clones was greatly simplified by their galactose dependence, which implied that the genes causing the phenotypes were P\textsubscript{GAL1} driven, thus effectively mapping their positions on the insert fragments. A single sequencing primer based on sequences within P\textsubscript{GAL1} allowed identification of each of these clones and will be of general use for the identification of other galactose-dependent clones. From 13 independent positive clones, seven known genes and one previously unknown gene were isolated.

Of the seven known genes, one of the strongest activators isolated was STE4, which confirms previous work in other labs. Genes were isolated that displayed a thousandfold range of levels of activation. Among the genes identified in the screen were three transcription factors, OBFI, PHO2, and MCM1. MCM1 has been shown previously to be involved in transcription of pheromone-responsive genes and PHO2 are involved in a variety of cellular processes (23, 24) but had not previously been implicated in mating response. It is possible that the transcriptional activation of FUS1 is an indirect result of increased levels of a general transcription factor. MSNI is a multicopy suppressor of a defect in the SNFI protein kinase (25). The fact that the mating response pathway involves at least four protein kinases suggests that MSNI may be performing a function in the mating response screen that is analogous to its role in suppressing the defective SNFI kinase.

\textit{MAT1} is an example of a gene that was isolated based on the phenotype caused by its expression in an inappropriate cell type. \textit{MAT1} is normally transcribed only in cells of mating type \(\alpha\). Its ectopic expression in an \(\alpha\) strain (such as NY19, in which the screen was performed) results in simultaneous production of \(\alpha\) and \(\alpha\)-specific gene products (26) and presumably autocrine activation of the pheromone response pathway.

The final known gene identified in the screen was \textit{STE11}, one of the protein kinases in the mating pathway. The gene was isolated twice, and neither clone contained the full-length gene. Both plasmids contained inserts lacking ≈40% of the amino terminus of the coding region. Interestingly, these truncated \textit{STE11} clones were potent activators of the mating response pathway, showing the highest levels of FUS1 induction of any of the genes isolated in the screen. Thus the
mating response activation screen revealed yet another interesting feature of the regulatable overexpression library, the possibility of identifying dominant truncation alleles. The galactose-dependent growth arrest gene may be another example of such a dominant truncation allele, because the reading frame of that clone is open up to the fusion breakpoint. Such truncation alleles are not represented in a high-copy-number library because they lack functional promoter elements.

The results of the library characterization, then, suggest that this regulatable yeast genomic DNA library will allow access to a number of different classes of genes that have phenotypic consequences when overexpressed but have been difficult or impossible to isolate to date using traditional high-copy-number libraries.

We thank P. Foreman and R. Fuller for critical comments on the manuscript. We would also like to thank P. Foreman, R. Fuller, and M. Schena for helpful discussions and suggestions and R. Fuller, R. Rugieri, R. Esposito, P. Webster, and B. Cairns for gifts of strains and oligonucleotides. This work was supported by National Institutes of Health Grants GM28428 and R37HG00198 (to R.W.D.). S.W.R. is a predoctoral trainee supported in part by a Training Grant 2TRGM07599 from the National Institutes of Health. S.J.E. was a Senior Fellow of the American Cancer Society.