Efficient PCR-based gene disruption in *Saccharomyces* strains using intergenic primers

Robert J.D. Reid, Ivana Sunjevaric, Mehdi Kedacche*****, and Rodney Rothstein†

Department of Genetics and Development, Columbia University College of Physicians and Surgeons, New York, NY 10032-2704

* Present address: Division of Human Genetics, Children's Hospital Medical Center, Cincinnati, OH 45229

† Corresponding author: Phone (212) 305-1733, FAX (212) 923-2090, E-mail rothstein@cancercenter.columbia.edu

Key Words - gene disruption; yeast; *Saccharomyces cerevisiae*; deletion consortium;

Abstract

Gene disruptions are a vital tool for understanding *Saccharomyces cerevisiae* gene function. An arrayed library of gene disruption strains has been produced by a consortium of yeast labs, however its use is limited to a single genetic background. Since the yeast research community works with several different strain backgrounds, disruption libraries in other common laboratory strains are desirable. We have developed simple PCR-based methods that allow transfer of gene disruptions from the S288C-derived strain library into any *Saccharomyces* strain. One method transfers the unique sequence tags that flank each of the disrupted genes and replaces the kanamycin resistance marker with a recyclable *URA3* gene from *Kluyveromyces lactis*. All gene-specific PCR amplifications for this method are performed using a preexisting set of primers that are commercially available. We have also extended this PCR technique to develop a second general gene disruption method suitable for any transformable strain of *Saccharomyces*.

Introduction

PCR-based methods have been used to streamline gene disruption techniques in *Saccharomyces cerevisiae* by eliminating the need for cloning in *Escherichia coli*. Such methods typically use chimeric primers to amplify a selectable marker. The 5' ends of these primers include 35 - 60 nucleotides of homology to a yeast gene. Upon transformation the homologous DNA ends promote two recombination events that replace the gene of interest with the selectable marker (Baudin *et al.* 1993). However, this method is inefficient due to the relatively short homology regions and requires several micrograms of DNA for each transformation (Wach *et al.* 1994; Wach 1996).

Recently, a consortium of yeast labs has used genome sequence information to produce chimeric primers for deletion of every yeast open reading frame (ORF). This has resulted in production of an arrayed library with disruptions of approximately ninety-three percent of yeast ORFs. Each strain is marked with a unique 20-mer sequence that identifies the gene disruption (Shoemaker *et al.* 1996). These unique markers, or barcodes, can be used to identify gene disruptions in a population of strains that survive a challenge to growth (Winzeler *et al.* 1999). While this arrayed library is an invaluable resource to the yeast research community, there are significant limitations to having a disruption library in a single genetic background. Many yeast experiments rely on specific assays that are difficult to reconstruct in 6000 new strains. Since gene disruption is one of the most powerful tools for analysis of gene networks, improved gene disruption methods will continue to be useful.

We have developed two PCR-based gene disruption methods that use relatively long regions of homology to improve gene targeting efficiency. Homologous DNA is PCR-amplified in one step, then fused to a selectable marker via PCR in a second step. Both methods take advantage of a preexisting set of commercialy available

primers designed to amplify the intergenic regions between yeast ORFs. Both methods also include a selectable marker that can be deleted from the genome allowing its reuse for additional gene disruptions. The first method was designed to specifically transfer gene disruptions from the library of strains produced by the gene disruption consortium, including the unique identifying 'barcode' sequence tags. The second method is a more general procedure that does not transfer the barcode tags and thus does not require DNA from the disruption library.

Materials and Methods

Materials

Taq DNA polymerase and dNTPs were purchased from Roche Diagnostics GmbH (Mannheim, Germany). All PCR primers were purchased from Research Genetics (Huntsville, AL). 5-Fluoroorotic acid was purchased from American Bioanalytical (Natick, MA).

Standard media preparations and growth conditions were used to culture yeast strains (Sherman *et al.* 1986). W1588-4A and W1588-4C strains are *RAD5* derivatives of W303 (Thomas and Rothstein 1989, and see Table 1). BY4741 and BY4743 are haploid and diploid parent strains for gene disruptions produced by the deletion consortium (Brachmann *et al.* 1998). Gene disruption strains produced by the yeast deletion consortium were obtained from Research Genetics. *ilv1*, *top3*, *sgs1*, *rad55, rad52* and *pet117* disruptions were obtained as haploid strains. Disruptions of the essential genes *cdc45* and *rnr1* were obtained as heterozygous diploid strains.

Plasmid construction

pWJ1042. Plasmid pWJ1042 is a *CEN/ARS* plasmid containing the *K. lactis URA3* gene flanked by direct repeats. Direct repeats flanking the *K. lactis URA3* gene were made using primers d2-Kl and u2-Kl (Table 2) to amplify a 192 bp DNA fragment from the *K. lactis URA3* 5' untranslated region. The PCR product was cut at *Cla*I and *Apa*I restriction sites contained in the d2-Kl and u2-Kl primers, and the resulting 142 bp fragment was cloned into *Cla*I and *Apa*I sites 3' to the *K. lactis URA3* gene in plasmid pWJ716 (Erdeniz *et al.* 1997).

pWJ1075. Plasmid pWJ1075 contains the *K. lactis URA3* gene flanked by direct repeats but does not contain *CEN* or *ARS* sequences. Primers ClaI-KlURA3' and SacI-KlURA5' were used to PCR amplify a 1.3 Kbp fragment from plasmid

pWJ716 containing the *K. lactis URA3* gene. The fragment was cloned into the *Sac*I and *Cla*I sites in pRS303 (Sikorski and Hieter 1989). The direct repeat was inserted as described above for pWJ1042.

pWJ1077. Plasmid pWJ1077 is the same as pWJ1075 except that nine nucleotide sequence changes were introduced to mutate the A-box consensus in the *K. lactis URA3* ORF while maintaining the wild-type amino acid sequence. PCR using primers ClaI-KlURA3' and ARSless-reverse amplified a 955 bp DNA. PCR using primers SacI-KlURA5' and ARSless-forward amplified a 408 bp DNA. A second round of PCR was performed using these products as template and primers ClaI-KlURA3' and SacI-KlURA5 to generate a fusion sequence with a mutated *ARS* A-box.

PCR

PCR primers pairs to amplify intergenic regions were designed to have an annealing temperature of 52°C (Vishy Iyer and Pat Brown, personal communication). Amplification of intergenic regions from W303 genomic DNA was performed as follows. Approximately 2 ng genomic DNA was added to 20µl reactions containing 0.5 µM of the specific intergenic primers, 200 µM dNTP mix, and 1.5 units Taq polymerase. Amplification was performed in an MJResearch PTC100 thermal cycler using the following cycle conditions: 94°C for 3 minutes followed by 30 cycles of 94°C for 30 seconds, 52°C for 30 seconds, 72°C for 1 minute and finally 72°C for 5 minutes. *K. lactis URA3* fragments were amplified from plasmid pWJ1042, pWJ1075 or pWJ1077 with the above conditions except the annealing temperature was 55°C. Genomic DNA for use as a PCR template was prepared by standard methods (Hoffman and Winston 1987).

PCR-mediated fusions were performed in 50 µl reactions containing 0.5 µl of PCR product for each template DNA (20 to 50 ng total), 0.5 µM primers, 200 µM dNTP

mix and 3.8 units of Taq polymerase. Cycle conditions were 94°C followed by 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1.5 or 2 minutes and finally 72°C for 10 minutes.

LiOAc transformation

Yeast transformations were performed essentially as described (Schiestl and Gietz 1989). Approximately 300 ng of each fusion DNA fragment was added to competent cells using conditions that give 10^6 transformants/µg circular plasmid DNA.

DNA sequencing

Upstream barcodes from deletion strains were amplified from genomic DNA using primers U1 and RGseqR to produce a 346 nucleotide PCR product. Upstream barcodes from the following strains were amplified for sequencing: YLR234W, YMR190C, YER058W, YER086W, YDR076W, YML032C, YLR103C, YER070W, YLR443W, YLR426W, YNR058W, YLR446W, YKR104W, YKR087C, YLR449W, YLR110C, YJR128W, YML009C, YJR079W, YLR432W, YNR068C, YNR059W, YKR098C. The upstream PCR product was isolated on an agarose gel and purified using the QIAquick PCR purification kit (QIAGEN, Valencia, CA). DNA sequencing was performed using the dideoxy chain termination method on an ABI377 sequencer (Applied Biosystems, Foster City, CA). Downstream barcodes were amplified using primers RGseqF and D1 to produce a 160 bp DNA for sequencing. Downstream barcodes were amplified from the following strains: YLR234W, YMR190C, YER058W, YER086W, YDR076W, YLR103C, YER070W, YLR443W, YLR426W, YNR058W. New deletion strains were sequenced after popout of the URA3 marker by amplifying a 255 bp DNA using primers D1 and U1.

Results

Transferring a gene disruption to a new strain

Our first method is designed to transfer a gene disruption from the library of gene disruption strains into a new genetic background. The specific goal is to transfer the barcode markers and provide a recyclable genetic marker in the process.

The deletion consortium strains were constructed using a chimeric primer PCR technique in which the kanamycin resistance selectable marker (kanMX4) is PCR amplified and attached to 45 bp homology from each side of the targeted coding region. The 45 bp of homology on the 5' side of the ORF is directly adjacent to and includes the start codon. Likewise, the 3' 45 bp of homology is adjacent to, and includes the stop codon thus generating precise disruptions of each ORF. Additionally, the long primers contain two unique sequence tags on the 5' ("UPTAG") and 3' ("DOWNTAG") ends of each disruption cassette providing the two unique identifiers for each gene disruption. As an example, the structure of the *TOP3* locus is shown in Figure 1A along with the structure of the *top3* deletion strain produced by the consortium. The UPTAG (checkered box) and DOWNTAG (black box) are flanked by standard primer binding sites for amplification of these tags.

For our method, the intergenic regions flanking the gene disruption are amplified using primers C and D for the intergenic regions (see Figure 2) and the common U2 and D2 primers flanking the UPTAG and DOWNTAG respectively (Figure 1B). The orthologous *URA3* gene from *Kluyveromyces lactis* complements *S. cerevisiae URA3* and is used as a selectable/counterselectable marker (Shuster *et al.* 1987). *URA3* is amplified in two separate PCR reactions to produce overlapping fragments using the internal kli3' or kli5' primers paired with the u2 or d2 primers respectively (Figure 1C). The u2 and d2 *URA3* primers contain 5' sequences that are the reverse and complement of the common primers U2 and D2 used to amplify the intergenic

DNAs. We refer to these primers with complementary 5' ends as adaptamers since they can be used in a generic manner to fuse DNA sequences by PCR (Erdeniz *et al.* 1997, see Table 2 for sequences). PCR fusion occurs when sequences containing these reverse and complementary ends are mixed and act as long primers in a round of PCR. Terminal primers are also added to these reactions to amplify the fused product (C with kli3 and D with kli5 in Figure 1D).

Gene disruptions are achieved by co-transformation of the two PCR products into yeast by the LiOAc procedure (Schiestl and Gietz 1989). Upon transformation, the split *URA3* marker recombines to generate a functional gene (Fairhead *et al.* 1996). Two additional recombination events with the chromosome replace the genomic locus (Figure 1E). The *K. lactis URA3* nucleotide sequence is only 71% identical to the *S. cerevisiae URA3* gene (Shuster *et al.* 1987). This low level of homology inhibits recombination between the *K. lactis URA3* sequence and the *S. cerevisiae URA3* genomic locus (Bailis and Rothstein 1990; Priebe *et al.* 1994). Furthermore, neither individual *URA3* fragment is sufficient to complement a *URA3* mutation, so rare random insertion events will not result in uracil prototrophy (Schiestl *et al.* 1993).

The *K. lactis URA3* plasmid used in this study contains a 143 bp direct repeat flanking the *URA3* coding region indicated by hatched boxes in Figure 1. Thus, after a successful gene disruption, "popout" recombinants can be selected using 5-FOA (Figure 1F). The unique sequence tags are maintained in these strains after popout so the strains can still be identified by automated methods (Winzeler *et al.* 1999).

Transfer of a gene disruption requires DNA from the original disruption strain and the gene-specific primers (labeled C and D in Figure 1A). All of the other primers and components can be applied to any gene disruption, making this a generic method. The gene-specific primers were chosen from a set of existing primers designed to amplify 6361 intergenic regions between every yeast ORF (Iyer *et al.* 2001). These

primers are commercially available (Research Genetics), and have common 5' sequence tags that are not homologuos to yeast genomic DNA sequences so that they can be used as adaptamers for PCR fusions (M. K. and R. R.,unpublished observations). Figure 2 illustrates intergenic primer binding sites for a centromere proximal portion of chromosome 4 (adaptamers are not to scale). Each "forward" intergenic adaptamer is homologous to the Watson strand and contains a nonhomologous 5' sequence tag labeled C (gray triangles in Figure 2, also see Table 2 for sequence). Likewise, each reverse intergenic adaptamer is homologous to the Crick DNA strand and contains a 5' nonhomologous sequence tag labeled D (Figure 2 and Table 2).

It is important to note that all the forward (C) and reverse (D) intergenic adaptamers are uniformly oriented with respect to the genetic map. However, the KanMX4 cassette in each disruption strain is oriented with respect to the start and stop codons and varies by gene orientation. Thus, for genes on the Watson strand, the left intergenic DNA is amplified by the upstream C intergenic adaptamer and the common U2 primer while the right intergenic DNA is amplified by the downstream D intergenic adaptamer and the D2 primer as pictured for the *TOP3* gene in Figure 1. For genes on the Crick strand, the left intergenic region is amplified by the C adaptamer and the common D2 primer while the right intergenic region is amplified using the D adaptamer and the U2 primer.

In this study, eight gene disruptions have been moved from the deletion consortium strains into the W303 strain background using the method described above (Table 3). After transformation with 300 ng each of the left and right DNA fragments, from 7 to 283 transformants were recovered. *TOP3*, *SGS1*, *PET117*, *ILV1*, *RAD55* and *RAD52* gene disruptions were made in haploid strains, while disruptions of essential genes *CDC45* and *RNR1* were made in diploid strains.

PCR verification was performed on several transformants from each experiment using an internal *URA3* primer and a primer that binds in an ORF adjacent to the disruption to amplify genomic DNA from each transformant (as illustrated by gray arrows in Figure 1E). Table 3 shows the percentage of uracil prototroph colonies that resulted in gene disruptions as determined by PCR.

Phenotypic assays were used to asses null phenotypes where possible. *top3* mutant strains displayed slow growth compared to a wild-type strain. *pet117* mutants failed to grow on medium containing glycerol as the sole carbon source. *ilv1* mutants failed to grow on synthetic medium lacking isoleucine. *rad52* and *rad55* mutants died upon exposure to 20 kilorads γ-irradiation. Finally, *cdc45* and *rnr1* heterozygous diploids displayed 2:2 segregation of lethality in haploid spores.

Upstream and downstream barcode sequences were amplified from each of the disruption strains listed in Table 3 for sequence analysis. The same sequences were also amplified from the consortium strains in order to compare barcode sequences before and after transfer. A single mutation was identified in the upstream barcode in the new *cdc45* disruption strain (tcccatacgacaagttgaga -> **G**cccatacgacaagttgaga). Thus, even without the use of a proofreading polymerase, only 1 of 13 barcodes acquired a mutation during the transfer process.

Our sequence analysis revealed that the upstream barcode in the consortium *sgs1* strain has a single basepair deletion as compared to the published database (http://www-deletion.stanford.edu/cgi-bin/tag_sequences/tagsequence.cgi). To approximate the scope of such errors, upstream barcodes from 14 additional consortium strains and downstream barcodes from 3 additional strains were sequenced and compared to the database. In all, 6 of 24 barcodes analysed in the consortium strains harbored single basepair substitutions or deletions (Table 4).

Thus, yeast researchers should be wary of such descrepancies and confirm barcode sequences in their favorite strains when necessary.

General gene disruption method

The above method was designed to address the specific task of moving disruptions from the consortium strains into any desired yeast strain. A more general gene disruption scheme, involving only a simple modification was also developed to take advantage of the intergenic adaptamers.

Figure 3 illustrates the procedure to disrupt the *ILV1* ORF on the Watson strand of chromosome V. The left and right intergenic regions are amplified in two separate PCR reactions with the appropriate intergenic adaptamers (Figure 3A). *URA3* is amplified in two separate PCR reactions to produce two overlapping fragments as in Figure 1C. The *URA3* adaptamers are designed to fuse with the intergenic PCR products in a second round of PCR. The left intergenic DNA is fused to the 5' *URA3* DNA via the D adaptamer and the right intergenic DNA is fused to the 3' *URA3* DNA via the C adaptamer (Figure 3B). The two fusion PCR products are co-transformed to produce the gene disruption. The targeted ORF is replaced by the *K. lactis URA3* gene flanked by direct repeats (Figure 3C). Direct repeat recombination leaves a small sequence in place of the original ORF (Figure 3D). This allows use of the *URA3* selectable marker in future experiments or gene disruptions.

DNA products obtained from the first rounds of PCR and subsequent fusion to *URA3* are shown for the *HIS4* and *ILV1* intergenic regions in Figure 4. Left and right intergenic regions were amplified using wild-type genomic DNA from strain W303 as a template (lanes 1 - 4). The *K. lactis URA3* 5' and 3' DNAs were amplified using plasmid pWJ1042 as a template (lanes 5 and 6). In this first round of PCR, 30 cycles of amplification typically yield 0.5 to 1 µg of product in a 20µl reaction.

DNA fusion by PCR was performed by mixing 10 to 25 ng of intergenic DNA with 10 to 25 ng of a *K. lactis URA3* PCR product in a 50 µl reaction. Typically, the template DNAs for fusion were simply diluted about 100-fold from the first PCR reactions and did not require purification. Fusion PCR reactions (Figure 4, lanes 7 to 10) yielded the specific fusion products and only rarely produced DNAs of different sizes. Yields of 1 to 5 µg of DNA per 50 µl reaction are common in a 30 cycle amplification. The fusion DNAs were taken directly from the PCR amplification mixtures as a source of DNA for yeast transformations. Purification of PCR fusion products is only necessary in the rare cases where multiple bands are observed.

Six gene disruptions performed by the general method were analyzed for efficiency and are listed in Table 5. Gene disruptions were verified by PCR, and when possible by assaying for a null mutant phenotype. Two PCR amplifications were performed to verify each gene disruption. Internal *K. lactis URA3* primers were paired with primers that bind in the left and right genomic DNA to amplify across each recombination junction (gray arrows in Figure 1E, also see Table 2). *SLA1* gene disruptions were assayed for growth on glycerol at 37°C. *ILV1* gene disruptions were assayed on synthetic drop-out media. *TRI2* gene disruptions were assayed by their slow growth phenotype.

Partial ARS consensus in K.lactis URA3

The majority of transformants for each gene disruption method resulted in a precise deletion of the targeted ORF. However, in an attempt to increase the overall efficiency of the methods, we also analyzed transformants that did not contain gene disruptions. In 11 gene disruption experiments 37 of 42 transformants (88%) that did not result in gene disruption also showed unstable uracil prototrophy. These transformants exhibited heterogeneous colony sizes on media lacking uracil and robust growth in the presence of 5-FOA, suggesting a transient extra chromosomal

maintenance of the *URA3* gene. We also identified a perfect match to an *ARS* A-box consensus sequence in the *K. lactis URA3* ORF (Broach *et al.* 1983). We hypothesized that unstable prototrophy could be due to circularization and *ARS*-dependent replication of the transformed DNA in the absence of integration. The *K. lactis URA3* ORF was cloned into a plasmid lacking CEN and ARS sequences. This plasmid did not transform yeast indicating that the A-box alone was not sufficient to promote plasmid replication.

A functional *ARS* contains the 11bp A-box that is a binding site for the origin recognition complex and less conserved elements generally referred to as B-boxes that have an anti-bent DNA structure (Eckdahl and Anderson 1990; Bell and Stillman 1992; Marilley 2000). Since it is feasible that fused intergenic DNA sequences can provide the less conserved elements for *ARS* function, DNA sequence changes were introduced into the *K. lactis URA3* ORF to mutate 9 of 11 bases of the A-box consensus while maintaining wild-type amino acid sequence. For *ILV1*, *CHL1* and *TOF1* gene disruptions the mutated *URA3* did not change the numbers of unstable uracil prototrophs (4 compared to 3 out of 36 transformants). In contrast, a *TOF2* gene disruption using the original *URA3* gene resulted in 9 unstable prototrophs from 11 transformants while the A-box mutant resulted in only 1 unstable prototroph from 11 transformants. Both *TOF2* intergenic regions reside near A-box consensus sequences. Thus it is likely that the *TOF2* gene disruption cassette contains DNA sequences that complete a functional *ARS* when combined with the *K. lactis URA3* A-box. To avoid complications, we now use the mutated (*ars*-) *K. lactis URA3* in our gene disruptions.

Discussion

In the post-genomic era, it is important to have simple and efficient methods for genome manipulation. Gene disruptions are important for molecular and genetic analysis in yeast. The existing library of arrayed yeast gene disruption strains is an important tool for these analyses. However, researchers often tailor a strain for use in a specific experiment while the set of disruptions exists in only a single strain background. Thus, it will often be preferable to transfer gene disruptions to a complex genetic background rather than introduce complex assays into the library of over 6000 disruption strains.

We have developed PCR-based methods to accomplish two separate tasks. One is transfer of a "barcoded" gene disruption from the consortium library into any desired *Saccharomyces* strain. The other method is a general gene disruption method applicable to any transformable yeast strain. Each of these methods share several advantages compared to other PCR-based methods. First, the efficiency of gene disruption is significantly increased due to the use of long homologous regions. Therefore, less DNA is required per transformation compared to existing chimeric primer methods. Second, each method allows counter selection of the *URA3* selectable marker via direct repeat recombination, in effect permitting the recycling of that marker for future use. This allows one to make serial gene disruptions in a strain using the same selectable marker. Third, all gene-specific primers are commercially available and the remaining components are generic. Finally, the *URA3* selectable marker can be amplified in quantity and used as a resource for many separate experiments making automation feasible.

Although the two adaptamer-directed gene disruption methods share many advantages, they were designed for separate tasks. The main reason to transfer a gene disruption from the library of consortium strains is to access the identifying tags

marking each disruption strain. In this case, DNA from a library strain is amplified using intergenic adaptamers paired with primers that amplify through the “barcode” tags. Orientation of the barcodes varies with gene direction so these must be paired correctly with the left or right intergenic adaptamer for amplification. In contrast, the general adaptamer-directed gene disruption method does not transfer barcodes, and also does not require DNA from a consortium strain for amplification. Fusion to the intergenic adaptamers is based on uniformly oriented C and D sequence tags making any gene disruption a standard protocol.

For the majority of the experiments performed, gene disruption was efficient. However, some problems were encountered. Gene disruptions that cause a slow growth phenotype, such as *top3* and *tri2* were recovered with lower efficiency. In such cases where a gene disruption causes a selective disadvantage, it may be best to perform the disruption in a diploid strain.

In other cases where gene disruption efficiency was low, the majority of the false positives were unstable uracil prototrophs. The most likely explanation for this is that the isolated *ARS* A-box motif we identified in the *K. lactis URA3* ORF is complemented for origin function when fused to certain intergenic DNA sequences. In these cases extrachromosomal DNA can be maintained. Removing the *ARS* element from *K. lactis URA3* reduces the occurrence of unstable prototrophs.

The adaptamer-directed gene fusion methods used in this study are versatile and efficient. Along with the gene disruption protocols described above, we have developed allele replacement and epitope fusion protocols based on adaptamer technology that are also PCR-based and use recyclable selectable markers (Erdeniz *et al.* 1997; Lisby *et al.* 2001). This suite of methods offers a unique ability to alter the *S. cerevisiae* genome with a minimal investment of time and materials.

Acknowledgments

This work was supported by grants HG00193 (RJDR) and HG01620 (RR) from the National Institutes of Health.

References

Bailis, A. M. and R. Rothstein (1990). “A defect in mismatch repair in *Saccharomyces cerevisiae* stimulates ectopic recombination between homeologous genes by an excision repair dependent process.” Genetics **126**: 535-47.

Baudin, A., O. Ozier-Kalogeropoulos, A. Denouel, F. Lacroute and C. Cullin (1993). “A simple and efficient method for direct gene deletion in *Saccharomyces* cerevisiae.” Nucleic Acids Research **21**(14): 3329-30.

Bell, S. P. and B. Stillman (1992). “ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex.” Nature **357**(6374): 128-134.

Brachmann, C. B., A. Davies, G. J. Cost, E. Caputo, J. Li, P. Hieter and J. D. Boeke (1998). “Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications.” Yeast **14**(2): 115-32.

Broach, J. R., Y. Y. Li, J. Feldman, M. Jayaram, J. Abraham, K. A. Nasmyth and J. B. Hicks (1983). “Localization and sequence analysis of yeast origins of DNA replication.” Cold Spring Harbor Symposia on Quantitative Biology **2**: 1165-73.

Eckdahl, T. T. and J. N. Anderson (1990). “Conserved DNA structures in origins of replication.” Nucleic Acids Res **18**(6): 1609-12.

Erdeniz, N., U. H. Mortensen and R. Rothstein (1997). “Cloning-free PCR-based allele replacement methods.” Genome Res **7**(12): 1174-83.

Erdeniz, N. and R. Rothstein (2000). “Rsp5, a ubiquitin-protein ligase, is involved in degradation of the single-stranded-DNA binding protein *rfa1* in *Saccharomyces cerevisiae*.” Mol Cell Biol **20**(1): 224-32.

Fairhead, C., B. Llorente, F. Denis, M. Soler and B. Dujon (1996). “New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination.” Yeast **12**(14): 1439-57.

Hoffman, C. S. and F. Winston (1987). “A ten-minute DNA preparation efficiently releases autonomous plasmids for transformation of Escherichia coli.” Gene **57**: 267-272.

Iyer, V. R., C. E. Horak, C. S. Scafe, D. Botstein, M. Snyder and P. O. Brown (2001). “Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF.” Nature **409**(6819): 533-8.

Lisby, M., R. Rothstein and U. H. Mortensen (2001). “Rad52 forms DNA repair and recombination centers during S phase.” Proc Natl Acad Sci U S A **98**(15): 8276-82.

Marilley, M. (2000). “Structure-function relationships in replication origins of the yeast *Saccharomyces* cerevisiae: higher-order structural organization of DNA in regions flanking the ARS consensus sequence.” Mol Gen Genet **263**(5): 854-66.

Priebe, S. D., J. Westmoreland, T. Nilsson-Tillgren and M. A. Resnick (1994). “Induction of recombination between homologous and diverged DNAs by double-strand gaps and breaks and role of mismatch repair.” Molecular & Cellular Biology **14**(7): 4802-14.

Schiestl, R. H., M. Dominska and T. D. Petes (1993). “Transformation of *Saccharomyces cerevisiae* with nonhomologous DNA: illegitimate integration of transforming DNA into yeast chromosomes and *in vivo* ligation of transforming DNA to mitochondrial DNA sequences.” Molecular & Cellular Biology **13**: 2697-2705.

Schiestl, R. H. and R. D. Gietz (1989). “High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier.” Current Genetics **16**(5-6): 339-46.

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

Shoemaker, D. D., D. A. Lashkari, D. Morris, M. Mittmann and R. W. Davis (1996). “Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy.” Nat Genet **14**(4): 450-6.

Shuster, J. R., D. Moyer and B. Irvine (1987). “Sequence of the *Kluyveromyces lactis* URA3 gene.” Nucleic Acids Research **15**(20): 8573.

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**(1): 19-27.

Thomas, B. J. and R. Rothstein (1989). “The genetic control of direct-repeat recombination in *Saccharomyces*: the effect of *rad52* and *rad1* on mitotic recombination at *GAL10*, a transcriptionally regulated gene.” Genetics **123**: 725-38.

Wach, A. (1996). “PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in S. cerevisiae.” Yeast **12**(3): 259-65.

Wach, A., A. Brachat, R. Pohlmann and P. Philippsen (1994). “New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces* cerevisiae.” Yeast **10**(13): 1793-808.

Winzeler, E. A., D. D. Shoemaker, *et al.* (1999). “Functional characterization of the S. *cerevisiae* genome by gene deletion and parallel analysis.” Science **285**(5429): 901-6.

Table 1. Yeast strains.

Strain	Genotype	Reference
W1588-4C	*MAT**a** ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5*	(Erdeniz and Rothstein 2000)
W1588-4A	*MAT**a** ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 RAD5*	(Erdeniz and Rothstein 2000)
BY4741	*MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	(Brachmann *et al.* 1998)
BY4743	*MAT**a**/MAT**a** his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 met15Δ0/MET15 lys2Δ0/LYS2 ura3Δ0/ura3Δ0*	(Brachmann *et al.* 1998)

W1588 strains are *RAD5* derivatives of W303 (Thomas and Rothstein 1989). BY strains are the parent strains of all deletion consortium strains obtained from Research Genetics (Huntsville, AL).

Table 2. PCR primers and adaptamers.

Primer Sequence	Name	Description
ccgctgctaggcgcgccgtg...	C	5' nonhomologous sequence tag used for all forward intergenic adaptamers.
gcagggatgcggccgctgac...	D	5' nonhomologous sequence tag used for all reverse intergenic adaptamers.
CTTGACGTTCGTTCGACTGATGAGC	kli-5'	*K. lactis URA3* internal 5' primer
GAGCAATGAACCCAATAACGAAATC	kli-3'	*K. lactis URA3* internal 3' primer
gtcagcggccgcatccctgcCCTCACTAAAGGGAACAAAAGCTG	d-Kl	3' *K. lactis URA3* "d" adaptamer. Nonhomologous region is the reverse and complement of D.
cacggcgcgcctagcagcggTAACGCCAGGGTTTTCCCAGTCAC	c-Kl	5' *K. lactis URA3* "c" adaptamer. Nonhomologous region is the reverse and complement of C.
GTCGACCTGCAGCGTACG	U2	5' primer for amplification of intergenic DNA from yeast deletion strains.
CGAGCTCGAATTCATCGAT	D2	3' primer for amplification of intergenic DNA from yeast deletion strains.
cgtacgctgcaggtcgacgggcccGTGTCACCATGAACGACAATTC	u2-Kl*	*K. lactis URA3* 5' adaptamer. The nonhomologous sequence is the reverse and complement of U2.
atcgatgaattcgagctcgatcgatGTGATTCTGGGTAGAAGATC	d2-Kl†	*K. lactis URA3* 3' adaptamer. The nonhomologous sequence is the reverse and complement of D2.
cagtctcagcaaTTTCGATGCAACCGGACTTGC	ARSless-forward#	Forward mutagenic primer used to alter *ARS* A-box
catcgaaattgctgagactgATGGATGAAAAGAAGACCAATTTGTGTGC	ARSless-reverse#	Reverse mutagenic primer used to alter *ARS* A-box
GATGTCCACGAGGTCTCT	U1	Primer used for sequencing barcodes.
CGGTGTCGGTCTCGTAG	D1	Primer used for sequencing barcodes.
GGGACAATTCAACGCGTC	RGseqR	Primer used for sequencing barcodes.
GACATCATCTGCCCAGATG	RGseqF	Primer used for sequencing barcodes.

All sequences are listed in the 5' to 3' direction. Lowercase sequences denote nonhomologous 5' segments or sequence tags on adaptamers respectively.
* The underlined sequence is an *Apa*I restriction site.
† The underlined sequence is a *Cla*I restriction site.
# The underlined sequences alter the A-box consensus in the *K. lactis URA3* gene.

Table 3. Consortium gene disruptions transferred into W303.

Gene	Percent disruption	Number examined
TOP3	30%	10
SGS1	50%	10
PET117	100%	12
ILV1	80%	40
RAD55	66%	12
RAD52	70%	7
CDC45	71%	7
RNR1	42%	12

Transformants containing gene disruptions were identified by PCR using primers on the right or left side of the gene disruption and a *K. lactis URA3* internal primer (see Figure 1E).

Table 4. Mutations identified in the consortium UPTAG sequences.

Gene	ORF	Published UPTAG	Alterations identified
SGS1	YMR190C	ccatgatgtaaacgatccga	ccatgatgtaaac**Δ**atccga
BIO3	YNR058W	cccgtactagcatttaatcg	cccgtactagcatttaatc**Δ**
IMD3	YLR432W	atcagactgcctaatgggcg	atcagactgcctaatgggc**C**
	YNR068C	taggacgagtcactgcatcg	taggacgagtcactgcatc**C**
MNT4	YNR059W	ggatattgcctcacacatcg	gg**Δ**tattgcctcacacatcg
UBP11	YKR098C	atattctgagacacgccgcg	ata**Δ**tctgagacacgccgcg

Sequences are listed in the 5' to 3' direction. **Δ** indicates a missing base. Base substitutions are indicated by a bold uppercase letter.

Table 5. Gene disruptions performed by the general adaptamer method.

Gene	Percent disruption	Number examined
SLA1	62%	24
ILV1	89%	9
TRI1	93%	14
TRI2	18%	11
TOF1[*]	80%	20
TOF2[†]	81%	11

Correct gene disruptions were defined by PCR analysis as described in Table 3.

[*] *K. lactis URA3* was amplified from plasmid pWJ1077 (mutated A-box) for the *TOF2* gene disruption results shown.

[†] *TOF1* gene disruption results are combined from experiments using *K. lactis URA3* amplified from pWJ1075 (+ A-box) and pWJ1077 (- A-box). In all other gene disruptions, *K. lactis URA3* was amplified from plasmid pWJ1042 (+ A-box).

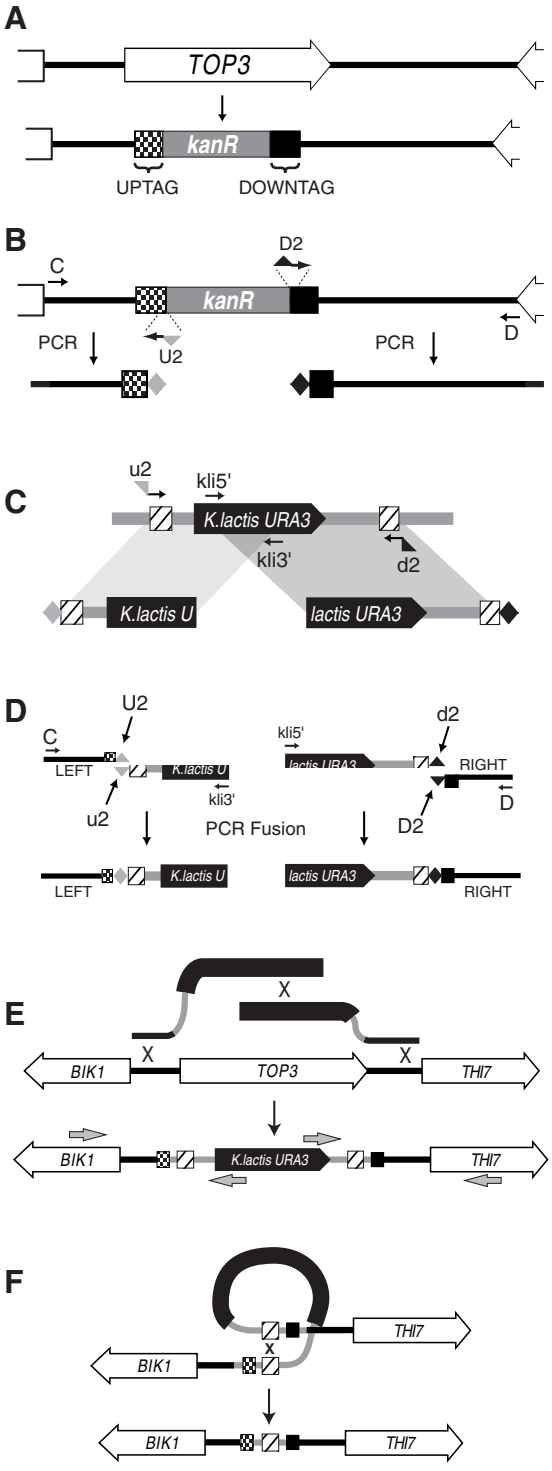
Figure Legends.

Figure 1. Amplification of "bar-coded" deletion cassettes and transfer of a gene disruption to a new strain. **A.** Structure of a *TOP3* gene disruption generated by the yeast deletion consortium is shown. White open arrows indicate open reading frames. The gray box represents the KanMX4 cassette. The UPTAG and DOWNTAG bar-codes are represented by a checkered box and a black box respectively. **B.** Adaptamers used to amplify the bar codes and the adjacent intergenic regions from a consortium gene disruption strain. C and D adaptamers are from the intergenic adaptamer set available from Research Genetics. Gray and black triangles indicate sequences used for fusion PCR in part D below while gray and black diamonds represent those sequences as duplex DNA. **C.** *K. lactis URA3* is amplified in two overlapping segments using adaptamer u2 with primer kli3' and adaptamer d2 with primer kli5'. Thick black arrow represents the *URA3* ORF, the gray line represents DNA from the pWJ1042 plasmid and hatched boxes represent the 143 base pair direct repeat sequences. **D.** Fusion occurs due to annealing of the complementary adaptamer tags in a second round of PCR (Diagonal arrows point to annealed sequences). Inclusion of the terminal primers (C and kli3**'** or kli5' and D) amplifies the fused product. **E.** Transformation of the two overlapping fusion PCR products and three recombination events, depicted as Xs, result in a *TOP3* gene disruption. The structure of the disruption includes the selectable marker, direct repeats flanking *URA3* and the UP and DOWN bar codes. Correct integration is confirmed by PCR using primers (gray arrows) that bind in adjacent ORFs paired with the internal *URA3* primers. **F.** A direct repeat recombination event, selected on 5-FOA medium, results in uracil auxotrophs so the *URA3* marker can be used for future disruptions in that strain.

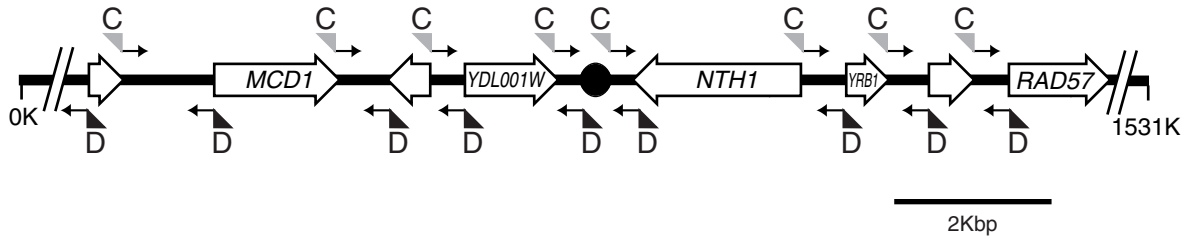
Figure 2. Orientation of adaptamers on yeast chromosome IV. A map of 15 kilobase pairs near the centromere of chromosome IV. The black circle represents the centromere. White arrows indicate known or predicted open reading frames. Adaptamers (not to scale) are shown as black arrows with gray (C) or black (D) triangles representing the generic C and D sequences.

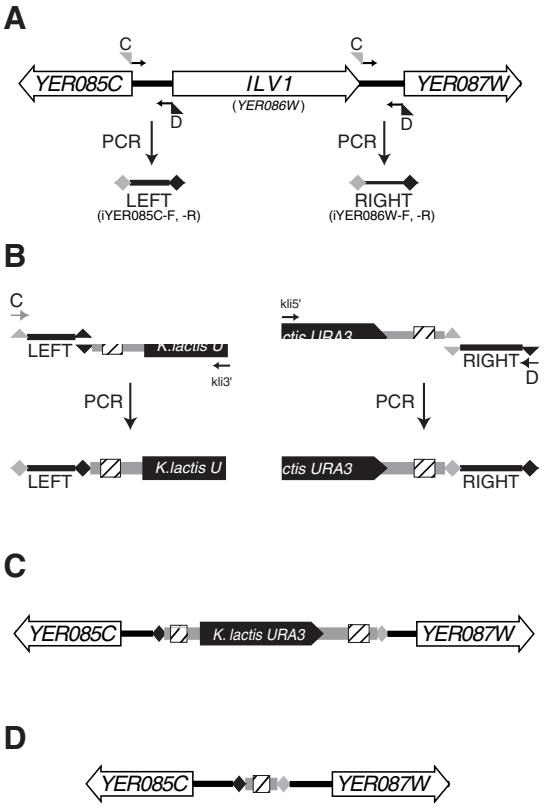
Figure 3. General method for adaptamer-directed gene disruptions. **A.** DNA sequences flanking an ORF to be deleted are amplified using intergenic adaptamers. Adaptamers are illustrated as in Figure 2. ORFs are labeled with standard gene names or systematic yeast ORF designations (also shown in parentheses). Vertical arrows indicate PCR amplification of the intergenic regions resulting in products containing the adaptamer tags. Intergenic adaptamers are named systematically with respect to the ORF immediately to the left on the genetic map. Primer designations are listed in parentheses below the PCR product. **B.** *K. lactis URA3* is amplified in two overlapping segments as illustrated in Figure 1B except that the fusions tags on the adaptamers are the reverse and complement of the C and D tags on the intergenic adaptamers. Fusion PCR is shown with outside primers C and kli3**'** or kli5' and D. Gene disruption occurs as diagramed in Figure 1E. **C.** Upon transformation, the two fusion DNA fragments recombine with genomic DNA resulting a gene disruption (as diagrammed in Figure 1E). The structure of the gene disruption includes the selectable marker and flanking direct repeats. **D.** Direct repeat recombination "popouts" of the *K. lactis URA3* marker as diagramed in Figure 1F can be selected on 5-FOA medium. The resulting genome structure is illustrated.

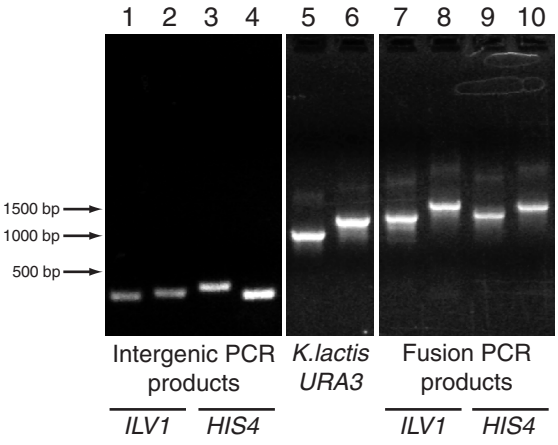
Figure 4. Examples of PCR amplification products for ILV1 (See figure 3) and HIS4 (not illustrated) gene disruptions. 2µl from each PCR reaction was analyzed by electrophoresis on a 0.8% agarose gel. Amplification of intergenic regions flanking the *ILV1* gene on chromosome V resulted in 352 and 380 bp left and right intergenic products (lanes 1 and 2). Amplification of intergenic regions flanking the *HIS4* gene on chromosome III result in 410 and 331 bp products (lanes 3 and 4). *K. lactis URA3* was amplified in two parts producing a 1000 bp 5' product and a 1263 bp 3' product (lanes 5 and 6). PCR fusion of the 1000 bp 5' *K. lactis URA3* sequence to the 352 bp *ILV1* left intergenic region produces a 1332 bp fusion product (lane 7). The 1263 bp *K. lactis URA3* 3' fused to the 380 bp *ILV1* right intergenic region produces a 1623 bp product (lane 8). The 5' *K. lactis URA3* DNA fused to the 410 bp *HIS4* left intergenic DNA produces a 1390 bp DNA (lane 9), and the 331 bp *HIS4* right intergenic fused to *K. lactis* 3' DNA produces a 1574 bp product (lane 10).


A
TOP3
kanR
UPTAG
DOWNTAG
B
C
D2
kanR
PCR
U2
PCR
D
C
u2
kli5'
K.lactis URA3
kli3'
d2
K.lactis U
lactis URA3
D
U2
C
LEFT
K.lactis U
kli3'
u2
PCR Fusion
kli5'
d2
RIGHT
D
D2
LEFT
K.lactis U
lactis URA3
RIGHT
E
X
X
X
BIK1
TOP3
THI7
BIK1
K.lactis URA3
THI7
F
THI7
X
BIK1
BIK1
THI7

Chromosome IV




A
C
C
YER085C
ILV1
(YER086W)
YER087W
D
D
PCR
PCR
LEFT
(iYER085C-F, -R)
RIGHT
(iYER086W-F, -R)
B
C
LEFT
K.lactis U
kli3'
kli5'
RIGHT
D
PCR
PCR
LEFT
K.lactis U
ctis URA3
RIGHT
C
YER085C
K. lactis URA3
YER087W
D
YER085C
YER087W


1 2 3 4 5 6 7 8 9 10
1500 bp
1000 bp
500 bp
Intergenic PCR products
K.lactis URA3
Fusion PCR products
ILV1 HIS4
ILV1 HIS4