A Novel Approach for Generating Full-length, High Coverage Allele Libraries for the Analysis of Protein Interactions

Phillip N. Gray‡, Katherine J. Busser, and Thomas G. Chappell

The yeast reverse two-hybrid method was developed to identify mutations disrupting protein-protein interactions. Adoption of the method has been slow, in large part, due to the high frequency of truncation and frameshift mutants typically observed with current protocols. We have developed a new strategy, based on in vitro recombinational cloning and full-length selection in *Escherichia coli*, to eliminate this background and dramatically increase the efficiency of the reverse two-hybrid protocol. The method was tested by generating an allele library of MyoD1 and selecting for alleles with defective interaction with Id1. Our results confirm that most of the interaction-defective alleles contain a single point mutation in the known interaction domain, the basic helix-loop-helix region. Moreover analysis of the crystal structure of MyoD reveals that the majority of these mutations occurred at the interaction interface. The results obtained using this novel approach for allele library generation demonstrate a significant advancement in the application of yeast reverse two-hybrid screens. Furthermore this method is applicable to any loss-of-function mutant screen where truncated proteins are a source of high background. *Molecular & Cellular Proteomics* 6:514–526, 2007.

As more researchers use systems approaches to study biological processes in metazoan cells, there is a need for more robust tools that mimic many of the techniques that have been available for years in model organisms. The ability to overexpress genome-wide collections of human genes, systematically alter gene expression using RNA interference, and test protein-protein interactions using a two-hybrid method or co-expression has greatly expanded the toolsets. Yeast two-hybrid technology (1) has been used to create global protein-protein interaction maps in *Helicobacter pylori* (2), *Saccharomyces cerevisiae* (3, 4), *Caenorhabditis elegans* (5), and *Drosophila melanogaster* (6, 7). More recently, the technology was used to produce two preliminary human interactomes investigating the potential protein-protein interactions of millions of human protein pairs (8, 9). These types of datasets are being mined by a large population of researchers for further characterization of specific interactions and pathways, mapping functional and interaction domains by screening allele variants of wild type ORFs. In the case of interaction domains, isolation of interaction-defective alleles (IDAs) from high coverage allele libraries allows the identification of residues and interfaces that mediate protein-protein or protein-nucleic acid interactions. IDAs may be isolated by a reverse two-hybrid method, which is a variation on a yeast two-hybrid method developed to identify cis and trans elements that disrupt protein interactions.

Current strategies for reverse two-hybrid analysis consist of three phases: 1) allele library generation, 2) counterselection, and 3) full-length selection. Allele libraries are produced by polymerase chain reaction followed by in vivo library assembly. When evaluating a protein-protein interaction using a counterselectable marker, such as *URA3* in the presence of 5-fluoro-orotic acid (5-FOA), a positive interaction will inhibit growth, whereas disrupted interactions will be resistant to 5-FOA (5-FOAR) (10–13). Both point mutations and truncated proteins may result in a disrupted interaction, but truncated proteins are usually not informative because large regions of the molecule may be absent. Therefore, isolating IDAs containing point mutations while selecting against truncated proteins is desirable. This can be achieved by incorporating a second step, positive selection, which requires the addition of an easily detected C-terminal fusion (10, 14) or the use of epitope tags (15). However, greater than 97% of 5-FOAR colonies (12, 14) are expected to contain alleles coding for truncated proteins. Thus, performing full-length selection after counterselection is not an effective solution to eliminate this high background. Separating the small percentage of full-length IDAs from background resulting from truncated proteins remains a challenge and represents a technical obstacle that has prevented widespread adoption of the technique.

We have developed a new strategy to address these issues...
that includes producing allele libraries in vitro and selecting for full-length proteins in Escherichia coli prior to analysis in yeast using Gateway technology. Gateway is a recombinational cloning technology based on λ phage recombination and facilitates the transfer of heterologous DNA sequences between vectors through site-specific attachment (attl) sites (16–19). We created a new Gateway vector, pDONR-Express, which facilitates the expression of ORFs as an N-terminal fusion to neomycin phosphotransferase and confers kanamycin resistance in E. coli. By selecting against truncated proteins in E. coli prior to IDA selection in yeast, almost all background normally associated with reverse two-hybrid screens is eliminated. Moreover when compared with gap repair-mediated library assembly, combining Gateway recombination with the efficiency of E. coli transformation allows for larger (105–107), more complex allele libraries to be evaluated.

Here we demonstrate the utility of pDONR-Express as a tool for generating full-length enriched allele libraries. We show that pDONR-Express is capable of selecting against truncated ORFs from allele libraries generated using mutagenic PCR. In addition, we generated an allele library of the basic helix-loop-helix (bHLH) protein MyoD1 (20, 21) and characterized its interaction with the HLH protein Id1, an interaction mediated by the HLH regions (22, 23). We determined that most interaction-defective alleles contained mutations within the bHLH region. Furthermore analysis of the crystal structure of the bHLH region of a MyoD homodimer (24) reveals that, not only are these mutations within the bHLH region, most are localized to one side of either helix 1 or helix 2 at the interaction interface.

MATERIALS AND METHODS

DNA Constructs—pDONR-Express was constructed using a traditional donor vector that contained the spectromycin resistance gene as the backbone. To express ORFs as entry clones in the Gateway system, a promoter was placed upstream of the attP1 site, and a single base pair change was made to remove a stop codon located 20 bp downstream of the 5’-end of attP1. Neomycin phosphotransferase (KanR) from pLenti3/V5 DEST (Invitrogen) was PCR-amplified to include E2F1 (GenBank accession number BC052160), LacZ (GenBank accession number L36850), the leucine zipper region of Fos (RefSeq accession number NM_005252), ATF2 (RefSeq accession number NM_001880.2), CDC3 (RefSeq accession number NM_007065.3), YWHAB (RefSeq accession number NM_003404), and the Ras association domain of RalGDS (GenBank accession number L07925).

Mutagenic PCR—The protocol was obtained from the Powers laboratory at University of California Davis.2 PCR conditions were set up to generate one mutation for every 60 bp using 100 ng each of the oligos 5’-GGGACCAAGTTTGTCAAAAAAGCAG-3’, 5’-GGGGACCCATTTGTACAAGAAAGCT-3’, or 5’-TAT ACC GGC TTT GGA ATC ACT-3’ and 5’-AGG CGA CAA CCT TGA TTG GAG AC-3‘; 5 μl of Taq Buffer without MgCl2; 15 μl of MgCl2 (50 mM); 4 μl of MnCl2 (5 mM); 1 μl each of 10 μm dGTP, dCTP, and dTTP and 0.5 μl of 10 μm dATP; 1 μl of Platinum Taq, and distilled H2O to 50 μl. Thirty cycles of PCR were performed at a Tm of 55 °C.

Allele Library Generation—The MyoD1 allele library was generated via PCR using 100 ng each of the oligos 5’-GGGGACCAAGTTTGTCAAAAAAGCAG-3’ and 5’-GGGGACCCATTTGTACAAGAAAGCT-3’ and pEXP22/MyoD1 (10 ng) as the template combined with 45 μl of Platinum Supermix HiFi (Invitrogen) with a Tm of 55 °C using standard PCR conditions (Note: some ORFs tend to be poor BP substrates and require additional flanking sequence outside the attB sites. This may be resolved by PCR amplifying with primers that anneal to the plasmid backbone at a distance of ~50–100 bp outside the attB sites.). PCR products were gel-purified using S.N.A.P. (Invitrogen) and quantified by measuring A260.

Library Transfer BP Reaction—The BP library transfer protocol was set up for a 1-kb ORF. The amount of PCR product may be scaled down for smaller ORFs. 450 ng of pDONR-Express, 200 ng of gel-purified PCR product (flanked by attB sites), 3 μl of BP Buffer, 8 μl of BP Clonase (Invitrogen), and TE to 20 μl were combined in one tube (Note: if using BP Clonase II, the following amounts should be added: 150 ng of pDONR-Express, 100 ng of gel-purified PCR product, 2 μl of BP Clonase II enzyme mixture, and TE to 10 μl). Reactions were incubated at room temperature (25 °C) for 20 h and then stopped by adding Proteinase K (1 μl/10 μl of reaction volume) and incubating at 37 °C for 10 min.

Kanamycin Titration—To determine the optimal kanamycin concentration for a particular ORF in the pDONR-Express system, two

---

2 T. Powers, personal communication.
transformations for a given BP reaction were set up (A and B). For each reaction, 1 μl of the BP reaction was transformed into 80 μl of TOP10 Electro-comp cells (Invitrogen) using a BTX ECM 630 electroporator (settings: 1700 V, 200 microhorns, 25 microfarads). Reaction A was recovered for 1 h in 1 ml of SOB + 1 mM IPTG at 37 °C at 250 rpm. Reaction B was recovered for 1 h in 1 ml of SOB at 37 °C at 250 rpm. Both reactions were serially diluted to 10⁻¹, and 100 μl of dilutions 10⁻², 10⁻³, and 10⁻⁴ were plated. Serial dilutions of transformation A were plated on LB/Spec (100 μg/ml) to test BP efficiency and LB/Kan at concentrations of 20, 30, 40, and 50 μg/ml + 1 mM IPTG. Serial dilutions of transformation B were plated on LB/Spec and LB/Kan (20, 30, 40, and 50 μg/ml). Plates were incubated at 30 °C for 24–36 h, and colonies were counted. An optimal kanamycin concentration will give you a maximum number of colonies under IPTG induction and a minimum number (or zero) without induction. It should be noted that the E. coli strains Mach1 (Invitrogen) and DH10B (Invitrogen) were also tested for pENTR-Express expression and produced poor results. Therefore we recommend using only TOP10.

pENTR-Express Allele Library Isolation—Once the optimal kanamycin concentration and target number of colonies were determined, the pENTR-Express library was transformed and plated to generate the desired number of clones for DNA isolation. 1 μl of the BP reaction was transformed into 80 μl of TOP10 Electro-comp cells as described above. Cells were recovered for 1 h in 1 ml of SOB + 1 mM IPTG at 37 °C at 250 rpm. Serial dilutions were performed and plated to the number of Kan⁺ colonies (Note: the transformation results obtained from the kanamycin titration step were used to estimate the colony-forming units/μl of BP reaction. This number was then used to estimate how much of the transformation should be plated on LB/Kan + 1 mM IPTG plates to get 20,000–30,000 colonies/plate.). Plates were incubated at 30 °C for 24–36 h. The remainder of the transformation was stored as a glycerol stock. After the transformant was determined, the glycerol stock was thawed on ice and plated out to a density of 20,000–30,000 colonies/plate on the appropriate number of LB/Kan (10–100 μg/ml) + 1 mM IPTG plates to produce the overall target number of Kan⁺ colonies. In addition, the glycerol stock was serially diluted and plated to check for loss in cell viability. Plates were incubated at 30 °C for 24–36 h, colonies were scraped, and DNA was isolated using S.N.A.P. MidiPrep (Invitrogen).

Library Transfer LR Reaction—Plasmid DNA was recovered from the library transfer BP reaction to yield allele libraries of the respective ORF as pENTR clones. 1 μg of pDEST22, 500 ng of pENTR-Express allele library, 3.5 μl of LR Buffer, 6 μl of LR Clonase, and TE to 20 μl were combined in one tube (Note: if using LR Clonase II, the following amounts should be added: 500 ng of pDEST22, 250 ng of pENTR-Express allele library, 2 μl of BP Clonase II enzyme mixture, and TE to 10 μl). The reaction was incubated at room temperature (25 °C) for 20 h and then stopped by adding Proteinase K (1 μl/10 μl of reaction volume) and incubating at 37 °C for 10 min.

pEXP22 Allele Library Isolation—The target number of clones from the LR reaction is the same number determined for the BP reaction. 1 μl of the LR reaction was transformed into 80 μl of TOP10 Electro-comp cells (Invitrogen) as described above. Cells were recovered for 1 h in 1 ml of SOB, version C at 37 °C at 250 rpm. Serial dilutions were performed and plated to the number of $[\text{3-AT}]$. Plates were incubated at 37 °C for 20–24 h. The remainder of the transformation was stored as a glycerol stock. After the transformant was determined, the glycerol stock was thawed on ice and plated out to a density of 20,000–30,000 colonies/plate on the appropriate number of LB/Amp (100 μg/ml) plates to produce the overall target number of Amp⁺ colonies. Plates were incubated at 37 °C for 20–24 h, colonies were scraped, and DNA was isolated using S.N.A.P. MidiPrep (Invitrogen).

Yeast Strains and Media—The reverse two-hybrid screen was conducted in the ProQuest yeast two-hybrid system (Invitrogen), which includes the S. cerevisiae strain MaV203 (MATα, leu2-3, 112, trp1-901, his3Δ200, ade2-101, gal4Δ, gal80A3, SPAL10-URA3, GAL1::lacZ, HIS3::GAL1::HIS3::LYS2, can1Δ, cyh2Δ). Complete synthetic medium dropout mixtures (BIO 101) were used to make media for all experiments. Synthetic complete medium (SC) containing 5-FOA was prepared as follows. 2× SC–Leu–Trp was prepared according to the manufacturer’s instructions, 5-FOA was added at either 0.05, 0.1, or 0.2%, the pH was adjusted to 4.5, and then the solution was filter-sterilized and combined with 2× agar cooled to 65 °C. SC–Leu–Trp–His + 3-amo-no-1,2,4-triazole (3-AT) was prepared by first preparing SC–Leu–Trp–His–3-AT and then autoclaving. Medium was cooled to 65 °C, 3-AT was added as a powder to a final concentration of either 10, 25, 50, or 100 μM and stirred until dissolved, and plates were poured.

Protocol for Conducting Screen—Yeast transformations were performed according to MaV203 competent yeast cell protocol (Invitrogen). Briefly 25 μl of cells were mixed with 1 μg of bait construct (pEXP32-Bait ORF) and 1 μg of prey allele library (pEXP22-Prey allele library). Next 180 μl of LiAc/polyethylene glycol solution was added, mixed, and incubated at 30 °C for 30 min. Then 10 μl of DMSO solution was filter-sterilized and combined with 2× agar cooled to 65 °C, and cells were heat shocked at 42 °C for 10 min. Cells were centrifuged at 1800 rpm, resuspended in 0.5–1 ml of distilled H₂O, and serially diluted to 10⁻², 100 μl of dilutions 10⁻¹ and 10⁻² were plated on SC–Leu–Trp, and 100 μl of undiluted cells and dilution 10⁻¹ were plated on SC–Leu–Trp + 5-FOA (Note: the entire transformation should be plated out to isolate as many mutants as possible.). Plates were incubated at 30 °C for 3–5 days. Then colonies were patched from SC–Leu–Trp–5-FOA onto SC–Leu–Trp (along with positive and negative control patches) and incubated at 30 °C for 2 days. Patches were replica-plate-plated onto SC–Leu–Trp and SC–Leu–Trp–His + 100 mM 3-AT, or the maximum 3-AT concentration at which the wild type interaction is observed. Plates were replica-cleaned until patches were no longer, or just barely, visible on the plate when held up to the light (typically after cleaning once or twice). Plates were incubated at 30 °C for 24 h, replica-cleaned again, and incubated at 30 °C until the positive control patch was clearly visible, typically 2–3 days, depending on the strength of the interaction and the concentration of 3-AT used. Yeast patches that did not display the wild type phenotype under histidine/3-AT selection were selected for plasmid isolation.

Plasmid Isolation from Yeast Using PureLink™—A 3 ml SC–Trp liquid culture was inoculated with cells from a fresh yeast patch from an SC–Leu–Trp plate. Yeast cells were collected from a 3-ml liquid culture by centrifugation in a tabletop centrifuge at 1,500 × g for 15 min. The cell pellet was resuspended in 240 μl of Resuspension Buffer containing RNase A, 10 μl of zymolyase (1.5 units/μl, Genotech catalog number 786-036) and 5 μl of β-mercaptoethanol were added, and cells were incubated at 37 °C for 30 min. Next 240 μl of Lysis Buffer were added, and contents were mixed gently by inverting the tube four to eight times. Tubes were incubated for 3–5 min at room temperature (Do not exceed 5 min.). Then 340 μl of Neutralization/Binding Buffer were added, and tubes were immediately mixed by gently inverting four to eight times and centrifuged in a tabletop centrifuge at maximum speed for 10 min to clarify the cell lysate. A PureLink spin column was placed inside a 2-ml collection tube. The supernatant was transferred into the spin column and centrifuged at room temperature at 10,000–14,000 × g for 30–60 s. The flow-through was discarded, and 650 μl of Wash Buffer were added to the column. The column was centrifuged at room temperature at 10,000–14,000 × g for 30–60 s, and the flow-through was discarded. The wash step was repeated followed by a dry spin to remove the residual wash buffer. The spin column was inserted into a clean 1.7-ml elution tube, and 70 μl of Elution Buffer or water were added to the center of
the column. The column was incubated at room temperature for 1 min and then centrifuged at maximum speed for 2 min. 5–10 μl of the purified DNA were used to transform TOP10 E. coli (Invitrogen). Transformants were plated out on medium containing ampicillin at 100 μg/ml to select for the pEXP22 backbone. Overnight cultures were grown, and plasmid DNA was isolated from E. coli transformants using the PureLink HQ kit plasmid DNA (Invitrogen). Plasmids were analyzed with the restriction enzyme BsrGI.

**Phenotype Confirmation**—All phenotypes were confirmed to verify that the initial mutant phenotypes were due to the isolated allele as opposed to a background mutation in the yeast. Following the transformation protocol outlined above, 26 individual transformations were set up using 100 ng of pEXP22/lid1 and 100 ng of each pEXP22/MycD1 allele. Transformations were plated onto SC–Leu–Trp plates and incubated for 3 days at 30 °C. A master plate was created by combining two to three individual colonies from each transformation and patching onto one SC–Leu–Trp plate with positive and negative control patches. The master plate was incubated overnight at 30 °C and then replica-plated onto SC–Leu–Trp–Ura and SC–Leu–Trp–His + 3-AT at concentrations of 10, 25, 50, and 100 μM. Plates were replica-cleaned until patches were barely visible on the plate when held up to the light (typically after cleaning once or twice), incubated at 30 °C for 24 h, replica-cleaned again, and incubated at 30 °C until positive control patches were clearly visible.

**Sequence Analysis**—For ORFs sequenced in the pEXP22 vector, sequencing reactions were performed using the oligos 5′-TAT ACC GGC TTT GGA ATC ACT-3′ and 5′-AGC GGA CAA CCT TGA TTG GAG AC-3′. For ORFs sequenced in the pENTR-Express vector, sequencing reactions were performed using the oligos 5′-GGA GAA CCT GCC TGC TGG TGC AAT CCA TCT-3′ and 5′-TAA TCA TGC GCA TAG TAT ATC GCC-3′. Other primers used for internal sequencing include: MyoD1, 5′-GCC TCG TGC TGG CAT GCC GCG CCC AG-3′; Cdc37a, 5′-GCC GAG TGC AGG AAA CT-3′; and Cdc37b, 5′-GAT CAA GCA CCT TGG CAT GC-3′. Sequences were analyzed using Sequencher. Translation of alleles and multiple sequence alignments were performed using Vector NTI.

**RESULTS**

**Design and Validation of the pDONR-Express Vector**—During the course of adapting yeast two-hybrid vectors for use with Gateway in vitro recombinational cloning, we were faced with the difficulty of designing different vectors for the forward and reverse applications of the technology. These difficulties could be circumvented, however, if the mutagenic protocols used for yeast reverse two-hybrid screens could be modified to eliminated nonsense and frameshift mutations prior to introducing allele libraries into yeast two-hybrid vectors. Because Gateway recombinational cloning can be used to move cDNA libraries from one vector backbone to another, we reasoned that we could use a similar approach to move ORF-specific allele libraries. In this way, we could first eliminate nonsense and frameshift mutations from our libraries using a C-terminal fusion as a read-through marker and then remove the C-terminal fusion as the libraries were moved to the yeast two-hybrid vector backbone.

Gateway cloning is based on the in vitro reconstitution of λ phage integration and excision reactions, which catalyze recombination between attB and attP sites (BP recombination) and attL and attR sites (LR recombination), respectively. Because of the size differences and distribution of stop codons among the four att sites, Gateway-adapted ORFs are expressed in vectors where they are flanked by inverted attB sites, which lack stop codons in the appropriate reading frames and allow read-through of the ORF at both the N- and C-terminal ends. ORFs are “stored” in entry vectors flanked by inverted attL sites that lack read-through capability on the N-terminal end. To achieve our design for allele library construction using standard Gateway vectors, the library would need to be shuttled through a series of expression and entry vectors, first to express the library with a C-terminal fusion in E. coli and then to express it with an N-terminal two-hybrid fusion in yeast (Supplemental Fig. 1). This approach, where both fusions are expressed with the allele library flanked by attB sites, required a total of four recombination reactions to create the library, eliminate nonsense and frameshift mutations, and finally utilize the library in a yeast reverse two-hybrid screen. Fig. 1 details a process where the four recombination reactions are reduced to two using a unique Gateway donor vector. Enablement of this approach required developing a vector capable of expressing fusion proteins where the ORF is flanked by inverted attL sites rather than attB sites.

pDONR-Express is a modified Gateway donor vector that was designed to express ORFs as a fusion to neomycin phosphotransferase. The key features that distinguish pDONR-Express from traditional donor vectors include (i) the EML promoter, a novel IPTG-inducible promoter, (ii) attP1*, a modified attP1 site, which contains an A to C mutation at position 20 that eliminates a stop codon and codes for an ORF that can be fused to a gene of interest, (iii) neomycin phosphotransferase (KanR), which is located downstream and in-frame with attP2, and (iv) lacI, which facilitates regulation of the EML promoter (Fig. 2). An inducible promoter was integrated into pDONR-Express to check the gene of interest for cryptic promoter activity, which will produce false positives by expressing partial ORFs fused to attL2-KanR. The vector may be used to select for ORFs coding for full-length proteins by simply inducing expression with IPTG after E. coli transformation and plating on medium containing kanamycin. The resulting fusion consists of attL1-ORF-attL2-KanR, and only alleles coding for full-length proteins will confer kanamycin resistance and produce colonies for DNA (i.e. allele library) isolation (Fig. 1, left). The entry clone allele library is then transferred to the two-hybrid AD vector through a second Gateway recombination reaction (LR reaction), yielding a full-length enriched expression library (flanked by attB sites) fused to Gal4-AD. As a result, clones lose the C-terminal fusion used for full-length selection and interactions may be evaluated in the original two-hybrid context (Fig. 1, right).

To test pDONR-Express for kanamycin selection and EML promoter induction, pDONR-Express was BP-crossed with five ORFs ranging in size from 300 bp to 3 kb and transformed into E. coli by electroporation. The resulting entry clones were tested for their ability to confer kanamycin resistance in the presence and absence of 1 mM IPTG. Table I shows high
FIG. 1. Strategy for achieving full-length selection in a novel Gateway entry vector and transferring alleles to traditional Gateway destination vector for IDA isolation. Left, the allele library is generated by PCR and BP-crossed into pDONR-Express. The resulting entry clones are expressed as N-terminal fusions to neomycin phosphotransferase, conferring kanamycin resistance in \textit{E. coli}. Right, following kanamycin selection, the full-length enriched allele library is LR-crossed into the yeast two-hybrid AD vector (pDEST22). The allele library is co-transformed into the yeast two-hybrid reporter strain with the DNA binding domain-bait plasmid. Wild type interactions grow on -Ura medium and are 5-FOA-sensitive. Mutants that fail to interact do not grow on -Ura medium and are 5-FOA-resistant.

FIG. 2. Vector map of pDONR-Express (top) and sequence of the promoter region (bottom). The EML promoter was created by inserting the lac operator (red) into the EM7 promoter. The modified attP1 site, attP1*, initiates translation at the first ATG and contains a mutation (TAG to TGC) to remove a premature stop codon. Also shown is the first 16 amino acids of the translated attP1* sequence.
numbers of kanamycin-resistant colonies in the presence of 1 mM IPTG for all ORFs tested, suggesting that the attL1-ORF-attL2-KanR fusion is being expressed. The absence of kanamycin-resistant colonies when IPTG is excluded indicates that expression of the fusion proteins is under the control of a functional lacIq gene product and lac operator within the EML promoter. The high number of colonies on LB/Spec plates verifies that all BP reactions were successful as non-reacted pDONR-Express contains the ccdB gene, which is toxic to TOP10 E. coli.

We found it necessary to titrate the amount of kanamycin used in the selection process for individual ORFs. A threshold concentration of kanamycin was found to exist for all ORFs evaluated where Kan− colonies appeared independently of IPTG induction. The background growth is most likely due to cryptic promoter activity and internal ribosome binding sites, which will produce a Kan− phenotype in the absence of a complete attL1-ORF-attL2-KanR fusion protein. To minimize this background, it was necessary to determine a kanamycin concentration that allows for a maximum number of colonies in the presence of IPTG while suppressing growth on kanamycin in the absence of IPTG. Of the ORFs tested, two (E2F1 and LacZ) produced colonies in the absence of 1 mM IPTG. However, the number of colonies on kanamycin medium lacking IPTG is minimal compared with the number on medium containing IPTG. For both of these ORFs, an average of 2–4% background was detected (Note: background is defined as the number of colonies on LB/Kan / number of colonies on LB/Kan + 1 mM IPTG).

Initial studies using ORFs with and without stop codons in the pDONR-Express system suggested that the presence of a stop codon would inhibit growth on medium containing kanamycin. To verify that the pDONR-Express system could discriminate alleles containing stop codons and frameshift mutations from those with missense mutations, an allele library of the leucine zipper region of Fos (258 bp) was generated by mutagenic PCR using conditions that generated one mutation per 60 base pairs. PCR products were BP-crossed into pDONR-Express, transformed into E. coli, and plated on LB/spectinomycin medium containing 1 mM IPTG. Several hundred colonies were patched onto LB/Kan + 1 mM IPTG, and plasmid DNA was isolated from clones displaying both Kan− and Kan+ phenotypes. Phenotypes were confirmed by retransforming the entry clones back into E. coli followed by induced expression and kanamycin selection. Confirmed ORFs were LR-crossed into pDEST22 for sequence analysis. Sequences were obtained for 27 clones displaying a Kan− phenotype and 29 clones displaying a Kan+ phenotype. A multiple sequence alignment was generated with translated Fos alleles from Kan− and Kan+ clones. Fig. 3A shows a schematic of the multiple sequence alignment. With the exception of one clone, sequence analysis of Fos alleles exhibiting Kan+ phenotypes show attB1-Fos-attB2 are in-frame, containing only missense mutations. The reading frame is maintained between Gateway reactions, so attL1, Fos, and attL2-KanR are expected to be in-frame in pENTR-Express.

The exception contains a 13-base pair deletion localized near the 5′-end of the ORF. Sequence analysis of clones exhibiting Kan− phenotypes show the Fos alleles containing either one or two deletions, nonsense mutations, or both, which would result in entry clones expressing either partial fusions or out-of-frame proteins that would not contain neomycin phosphotransferase. These results suggest that pDONR-Express is capable of discrimination against truncated ORFs in the majority of cases. The 13-base pair deletion in the Kan− clone results in a frameshift mutation that generates two tandem GGA codons followed by a GGG, AGC, and TGA and would result in a frameshift to maintain the correct reading frame. GGA codons have been reported to be associated with non-programmed −1 frameshifting (26). Thus, we believe the ka-

### Table I

| ORF (size) | [Kanamycin] | Recovery in 1 mM IPTG | No. of colonies ($10^4$) |
|------------|-------------|------------------------|--------------------------|
|            | ug/ml   |                         | On LB/Spec | On LB/Kan + 1 mM IPTG | On LB/Kan |
| MyoD (1 kb) | 20      | +                       | 910         | 35                     | 0          |
|            |         |                         | 860         | 0                      |            |
| E2F1 (1.3 kb) | 30       | +                       | 234         | 50                     | 3.3        |
|            |         |                         | 220         |                         |            |
| RalGDS (300 bp) | 20       | +                       | 2100        | 45                     | 0          |
|            |         |                         | 1970        |                         |            |
| Fos (300 bp) | 50       | +                       | 1360        | 138                    | 0          |
|            |         |                         | 1330        |                         |            |
| LacZ (3 kb) | 50       | +                       | 786         | 245                    | 3.4        |
|            |         |                         | 921         |                         |            |
FIG. 3. Schematic of multiple sequence alignments of mutagenized open reading frames. Shown are protein sequence alignments of mutagenized ORFs for Fos (A), ATF2 (B), YWHAB (C), and CDC37 (D), distinguishing between growth and non-growth on kanamycin medium. Approximately 350 ORFs containing more than 2000 mutations were analyzed.
namycin-resistant phenotype displayed by this clone is the result of a −1 frameshift, which restores the appropriate reading frame for neomycin phosphotransferase expression.

In addition to Fos, heavily mutagenized allele libraries were generated for the ORFs ATF2 (627 bp), YWHAB (735 bp), and CDC37 (1131 bp). Kan− colonies were isolated as outlined above. Kan+ colonies were isolated by plating transformants directly onto LB/Kan + 1 μg IPTG. All clones were sequenced directly from the pENTR-Express vector. For the ATF2 (Fig. 3B) and YWHAB (Fig. 3C) ORFs, we observed results similar to those obtained with Fos. Of the 50 ATF2 Kan− clones sequenced, only three coded for truncated proteins (94% full length), and 49 of 49 Kan− clones coded for truncated proteins. Likewise 48 of 50 YWHAB Kan− clones coded for full-length proteins (96% full length), and 55 of 55 Kan− coded for truncated proteins. All three ATF2 Kan+ clones that coded for truncated proteins contained a single nonsense mutation within the first 100 bp of the ORF. For the two Kan− YWHAB clones that coded for truncated proteins, each contained either a single base pair deletion or a nonsense mutation within the first 100 bp of the ORF. Thus, for three of the ORFs evaluated (Fos, ATF2, and YWHAB), it appears that some clones containing nonsense or frameshift mutations in the 5′-region may be able to slip through the kanamycin selection. For CDC37 (Fig. 3D), the background was a little higher: 41 of 46 Kan+ clones coded for full-length proteins (89% full length), and 46 of 48 Kan− coded for truncated proteins. Four of the five Kan+ clones we isolated that coded for truncated proteins contained a single base pair deletion that was not biased toward the 5′-end of the ORF. These mutants required a +1 frameshift to maintain the proper reading frame for kanamycin resistance, and three of the four contained an AGG codon immediately preceding the deleted nucleotide (the fourth deletion mutant contained an AGA), which has been reported to be associated with +1 programmed frameshifts (26). Thus, we believe the kanamycin-resistant phenotype displayed by these clones is the result of a +1 frameshift, which restores the appropriate reading frame for neomycin phosphotransferase expression. The remaining clone contained a nonsense mutation immediately followed by an ATG that was not biased toward the 5′-end. In addition, CDC37 was the only ORF observed to have in-frame Kan− clones. Both clones possessed unique mutations, so we can only speculate that these mutations are interfering with the function or stability of the fusion protein.

Allele Library Generation—The generation of an allele library requires a minimum number of clones to be isolated for good library representation. This target number of clones/collection will depend on the size of the ORF under study with larger ORFs requiring a higher target number. Errors generated by Taq polymerase are reported to occur in a biased manner (i.e., not all types of nucleotide changes occur at equal frequencies). As a result, the number of mutations per DNA sequence generated during PCR is not expected to follow the Poisson distribution (27, 28). In an effort to create general guidelines, we reasoned that for a 1-kb ORF, which possesses ~333 codons, ~1000–2000 333, (or 333,000–666,000) clones would be sufficient to generate good library representation. In creating the allele library for MyoD1, we decided that a minimum of 500,000 individual Kan+ clones was sufficient to provide good library representation for the 1081-bp ORF. This target number of colonies was exceeded with ~700,000 Kan+ colonies produced. The resulting pENTR library was isolated and LR-crossed into pDEST22. The target number of colonies (Amp+) from the LR reaction was 500,000. Approximately 2,600,000 Amp+ colonies were produced, and the resulting pEXP22-MyoD1 allele library was isolated.

Reverse Two-hybrid Analysis of the Id1-MyoD1 Interaction—The ProQuest (Invitrogen) yeast two-hybrid system was used to analyze the Id1-MyoD1 interaction. The pEXP22-MyoD1 allele library was co-transformed with pEXP32-Id1 into MaV203, which contains the SPAL10::URA3 reporter gene. Activation of this reporter by a protein-protein interaction converts 5-FOA into the toxic product 5-fluorouracil, which inhibits yeast growth. Thus, interaction-defective alleles may be selected out of libraries consisting largely of wild type alleles (12, 13). Interaction-defective alleles of MyoD1 were selected on medium containing 5-FOA at concentrations of 0.05, 0.1, and 0.2%. Approximately 1% of 10,000 transformants displayed strong 5-FOA5 phenotypes, most of which were observed on media containing 0.1 and 0.2% 5-FOA. 87 5-FOA5 clones, plus positive (Id1-MyoD1) and negative (Id1-RalGDS) controls, were tested for their ability to activate the HIS3 reporter in the presence of 3-AT, an inhibitor of His3p, at concentrations of 10, 25, 50, and 100 μM. 5-FOA5 clones that behave identically to wild type under histidine/3-AT selection may contain a mutation in the URA3 reporter gene rather than a mutant MyoD1 allele. Thus, this second step positive selection may serve to separate 5-FOA5 strains containing true mutants from those harboring wild type.

Sequence data were obtained from 32 MyoD1 alleles displaying the 5-FOA5 phenotype and suppressed growth on histidine-deficient medium supplemented with 3-AT. Of the 32 clones, 15 were wild type, 14 contained a single missense mutation, one contained three missense mutations, one contained a point mutation in the leader sequence, and one contained a truncated ORF. Sequences of the 15 alleles containing missense mutations within the MyoD1 ORF were translated and aligned with a MyoD1 template sequence using ClustalW (not shown). With the exception of clone 12 (which contains three point mutations), all alleles possess a single point mutation in either helix 1 or helix 2 within the bHLH domain (two of 15 contain a single point mutation in helix 1, and 12 of 15 contain a single point mutation in helix 2). The positions of these mutations in the bHLH region is diagrammed in Fig. 4, bottom. To confirm the initial mutant phenotypes, plasmid DNA from 16 mutant alleles (the truncated mutant was not included) and 10 wild type clones was co-
transformed into MaV203 with pEXP32-Id1. Transformants were tested for their ability to activate the *URA3* reporter as well as the *HIS3* reporter in the presence of 10, 25, 50, or 100 mM 3-AT (Fig. 4, top). The 3-AT titration provides information on how a particular mutation affects the interaction. Mutations that completely disrupt the interaction are unable to grow in the presence of low concentrations of 3-AT (10 mM), whereas mutations that weaken the interaction can survive on higher levels (25–100 mM). Of the 10 wild type clones, eight (clones 3, 10, 15, 16, 20, 21, 25, and 26) produced strong *URA*/*H11001* and *HIS3*/100 mM 3-AT phenotypes, whereas two (clones 11 and 13) displayed minimal growth under these conditions.

**Fig. 4. MyoD1 alleles.** Top, phenotype confirmation. Individual transformants containing pEXP32-Id1 and either wild type or mutant MyoD1 alleles (pEXP22-MyoD1) were tested for their ability to activate the *URA3* reporter as well as the *HIS3* reporter in the presence of 10, 25, 50, or 100 mM 3-AT. + = positive control (Id1/MyoD1), − = negative control (Id1/RalGDS). Bottom, diagram of the locations of mutations within the bHLH region. With the exception of one clone, all IDAs contained a point mutation in the bHLH region, the known interaction domain. −*LW*, without Leu and Trp; −*LWH*, without Leu, Trp, and His; −*LWU*, without Leu, Trp, and Ura.
TABLE II

Summary of MyoD1 alleles containing point mutations

Shown is a summary of MyoD1 mutant alleles containing point mutations and their phenotypes under histidine/3-AT selection. The table lists all amino acid changes from alleles containing point mutations. The [3-AT] listed is the concentration required to inhibit growth under histidine selection.

| Mutation | Clone/allele | 3-AT phenotype |
|----------|--------------|----------------|
| T115A    | 12           | 100            |
| F129S    | 1            | 10             |
| L132P    | 14           | 10             |
| K146T    | 24           | 25             |
| V147M    | 7            | 50             |
| V147A    | 19           | 50             |
| L150R    | 9            | 10             |
| R151H    | 12           | 100            |
| R151C    | 6            | 100            |
| I154T    | 8, 18        | 10             |
| E158K    | 2            | 10             |
| L160P    | 4, 17, 22    | 10             |
| L164P    | 23           | >100           |

conditions (Fig. 4, top). The reason for this observation is unclear. These clones may contain mutations in their promoters, decreasing the expression of wild type MyoD1. All mutant alleles (clones 1–2, 4–9, 12–14, 17–19, 22–24) were unable to activate the URA3 reporter as indicated by the absence of growth on SC-Leu-Trp-Ura plates and displayed varying sensitivities to 3-AT (Fig. 4, top). Table II lists a summary of the MyoD1 alleles and the maximum [3-AT] required to suppress growth. Clone 23 (L164P) was the only mutant displaying a strong growth phenotype in the presence of 100 mM 3-AT.

Analysis of MyoD Crystal Structure—To validate our results, we used the crystal structure of MyoD bHLH-DNA complex (Protein Data Bank code 1MDY) as a model (24). In this structure, the bHLH domain of MyoD (containing a C135S mutation in the loop) is complexed with a synthetic strand of DNA as a homodimer. Fig. 5 (left and right) shows the MyoD bHLH homodimer represented in solid ribbon with one strand colored red and the other colored blue. Residues within the bHLH region that were found to be mutated in interaction-defective alleles containing a single codon change are highlighted in yellow on the red strand. Most of these residues are located at the interaction interface and code for either aliphatic or aromatic amino acids, which have been reported to be common at binding surfaces (29). Moreover Fig. 5, left, shows that these mutations are located outside the DNA binding domain of the bHLH.

Table III lists a summary of residues that appear to facilitate interaction between the two molecules based on analysis of the crystal structure. The molecules interact in such a way that residues in helix 1 of the red strand interact with residues in helix 2 of the blue strand (Fig. 5). This is the case with all residues except Leu160 where both Leu160 residues are located in helix 2. Moreover four of six interactions can be found in both orientations. For example Phe129 of helix 1/strand A interacts with Leu150 of helix 2/strand B and vice versa (i.e. Leu150 of helix 2/strand A interacts with Phe129 of helix 1/strand B). This is the case for the Phe129-Leu150 and Leu132-Ile154 interactions (four total). The other interaction is Val147-Val125 where Val147 of helix 2/strand A interacts with Val125 of helix 1/strand B. We isolated alleles containing mutations in all of these positions (Val125, Phe129, Leu132, Val147, Leu150, Ile154, and Leu160) in the bHLH region except Val125.

Table III also lists the corresponding residues found in Id1 for both strands A and B. All residues are identical between Id1 and MyoD1 except at positions 125 and 129. However, the class of amino acid at these positions is conserved. MyoD1 contains a phenylalanine at position 129; Id1 contains a tyrosine; both are aromatic. MyoD1 contains a valine at position 125; Id1 contains a methionine; both are aliphatic. The level of conservation at these residues suggests that the Id1-MyoD1 complex should form a structure similar to that of the MyoD homodimer, so it is reasonable to model the interactions of the Id1-MyoD1 complex based on the 1MDY crystal structure.

We compared the phenotypes observed under histidine/3-AT selection with the location of the point mutation in the crystal structure of each allele and found a good correlation for alleles containing mutations at the interaction interface. Five of the seven alleles that contain mutations at the interaction interface (i.e. F129S, L132P, L150R, I154T, and L160P) failed to grow under histidine selection in the presence of 10 mM 3-AT (Fig. 4, bottom, and Table II). These results suggest that the interaction between Id1 and these alleles is severely, or completely, disabled. The F129S and I154T mutations transition from aromatic and aliphatic to nucleophilic amino acids, which are not expected to interact with leucine. Likewise the L150R mutation transitions from an aliphatic to a basic residue and is not expected to interact with tyrosine (see Table III). The L132P mutation most likely disrupts helix 1, and the L160P mutation disrupts helix 2. In contrast, alleles containing the V147M or V147A mutations required 50 mM 3-AT to suppress growth, suggesting that these alleles still interact with Id1 but with reduced affinity. This is not surprising because the class of amino acid is conserved in the V147M mutation, both are aliphatic, and a transition from valine to alanine in the V147A mutation substitutes aliphatic for small.

Alleles containing mutations outside the interaction interface include K146T, R151C, E158K, and L164P. Ma et al. (24) report a hydrogen bond between Asn126 of helix 1 and Lys146 of helix 2 that is thought to stabilize the molecule. The K146T mutation changes the residue from basic to nucleophilic, which would destroy the hydrogen bond with Asn126 and destabilize the molecule. The allele containing this mutation required 25 mM 3-AT to suppress growth, suggesting a weakened interaction with Id1. The Arg151 and Glu158 residues are located in the bHLH region one position away from the interaction interface. The allele containing the R151C mutation...
required 50 mM 3-AT to suppress growth, suggesting that this allele still interacts with Id1 but with reduced affinity. The allele containing the E158K mutation failed to grow under histidine selection in the presence of 10 mM 3-AT, suggesting a disrupted interaction with Id1. These two residues are not conserved between Id1 and MyoD1; therefore the 1MDY crystal structure cannot be used as a model to determine the role these residues play in the interaction with Id1. These residues could stabilize the bHLH through intramolecular interactions with regions not included in the crystal structure. Allele 12 contains three point mutations (T115A, R151H, and N204D), with one located within helix 2 of the bHLH region (R151H), and displays a phenotype similar to that of allele 6, which contains a similar mutation (R151C). The Leu164 residue is within helix 2, facing away from the interaction interface, and alleles containing L164P behave similarly to wild type under histidine/3-AT selection. However, this mutation probably distorts helix 2, weakening interaction with Id1, because this allele is unable to activate the *URA3* reporter. Clone 5 was the only allele isolated with a mutation outside the MyoD1 ORF. This allele is unable to activate the *URA3* reporter and failed to grow under histidine selection in the presence of 100 mM 3-AT.

We have demonstrated the ability of pDONR-Express to select against truncated proteins from four different allele libraries generated through mutagenic PCR. Our results suggest that greater than 90% of alleles generated using this system should code for full-length proteins. The percentage of full-length protein was the lowest for the largest ORF; however, the fact that most of these deletion mutants contained a conserved codon immediately upstream of the deleted base pair that is associated with programmed +1 frame-shifts suggests that the sequence of the ORF may be more of a contributing factor to the background growth rather than the size.

We used pDONR-Express to generate a full-length enriched MyoD1 allele library and selected for interaction-defective alleles. 15 of 18 interaction-defective alleles contained a single point mutation in the known interaction domain. Thus, this system is capable of identifying interaction domains within an ORF; this is significant when no structure data are available. Fortunately for us, the three-dimensional structure of the bHLH of MyoD had been solved, so we were able to map the positions of the point mutations from the interaction-defective alleles onto the crystal structure. Of the 10 point mutations within the bHLH region six were located at the interaction interface at residues that appear to facilitate protein binding. In fact, a total of seven residues appear to mediate protein binding between the two molecules, and we isolated interaction-defective alleles containing mutations at six of these seven positions.

The data obtained from our reverse two-hybrid analysis of Id1-MyoD1 demonstrate the potential of our new strategy for generating allele libraries for reverse two-hybrid analysis of protein interactions. This strategy has several advantages over existing methods. First, generating allele libraries in *vitro* with Gateway technology is more efficient than gap repair, which may result in 9% of plasmids without insert (10). Sec-

---

**DISCUSSION**

We have demonstrated the ability of pDONR-Express to select against truncated proteins from four different allele libraries generated through mutagenic PCR. Our results suggest that greater than 90% of alleles generated using this system should code for full-length proteins. The percentage of full-length protein was the lowest for the largest ORF; however, the fact that most of these deletion mutants contained a conserved codon immediately upstream of the deleted base pair that is associated with programmed +1 frame-shifts suggests that the sequence of the ORF may be more of a contributing factor to the background growth rather than the size.

We used pDONR-Express to generate a full-length enriched MyoD1 allele library and selected for interaction-defective alleles. 15 of 18 interaction-defective alleles contained a single point mutation in the known interaction domain. Thus, this system is capable of identifying interaction domains within an ORF; this is significant when no structure data are available. Fortunately for us, the three-dimensional structure of the bHLH of MyoD had been solved, so we were able to map the positions of the point mutations from the interaction-defective alleles onto the crystal structure. Of the 10 point mutations within the bHLH region six were located at the interaction interface at residues that appear to facilitate protein binding. In fact, a total of seven residues appear to mediate protein binding between the two molecules, and we isolated interaction-defective alleles containing mutations at six of these seven positions.

The data obtained from our reverse two-hybrid analysis of Id1-MyoD1 demonstrate the potential of our new strategy for generating allele libraries for reverse two-hybrid analysis of protein interactions. This strategy has several advantages over existing methods. First, generating allele libraries in *vitro* with Gateway technology is more efficient than gap repair, which may result in 9% of plasmids without insert (10). Sec-

---

**TABLE III**

Summary of the putative interactions between residues on each bHLH molecule (red and blue in Fig. 5) of MyoD and corresponding residues on Id1

| MyoD red | Id1 | MyoD blue | Id1 |
|----------|-----|-----------|-----|
| Helix 1  | Phe<sup>129</sup> | Phe<sup>129</sup> | Helix 2 | Leu<sup>150</sup> | Leu<sup>150</sup> |
|          | Leu<sup>132</sup> | Leu<sup>132</sup> |          | Ile<sup>154</sup> | Ile<sup>154</sup> |
| Helix 2  | Val<sup>147</sup> | Val<sup>147</sup> | Helix 1 | Val<sup>125</sup> | Met<sup>125</sup> |
|          | Leu<sup>150</sup> | Leu<sup>150</sup> |          | Phe<sup>129</sup> | Leu<sup>132</sup> |
|          | Ile<sup>154</sup> | Ile<sup>154</sup> |          | Leu<sup>160</sup> | Leu<sup>160</sup> |
| Helix 2  | Leu<sup>160</sup> | Leu<sup>160</sup> | Helix 2 | Leu<sup>160</sup> | Leu<sup>160</sup> |
on, the high transformation efficiencies of E. coli allow for larger, more complex allele libraries to be generated. Third, these high titer allele libraries are preselected to eliminate truncations and frameshifts, and each subsequent yeast transformation event is then used to perform a real reverse two-hybrid selection for an interaction-defective allele rather than current methods where nearly 97% of the yeast transformation events are wasted selecting for truncated proteins (10, 12). We have generated a few dozen independent alleles in each prey protein from three independent screens, picking and rescreening 50–100 5-FOAR colonies to eliminate the spontaneous mutants that lead to 5-FOA resistance. In contrast, for nine reverse two-hybrid screens, Endoh et al. (10) had to screen 12,724 5-FOAR isolates for green fluorescent protein staining to isolate 28 alleles of which 22 were independent. Thus, this new method has reduced the frequency of false positives in reverse two-hybrid screens –100-fold, and 10^5 yeast transformants are now sufficient to do comprehensive mutagenic coverage of an allele library of a single ORF.

Fourth, Gateway technology allows library transfer from the entry vector to the yeast two-hybrid expression vector. Thus, by separating full-length selection from reverse two-hybrid analysis, protein-protein interactions may be studied in the original two-hybrid context. In contrast, the Endoh et al. (10) system requires both an N- and C-terminal fusion. Although the authors claim that 95% of interactions were not affected by the presence of these two fusions, Huh et al. (30) have reported the potential mislocalization of nearly 40% of 4156 yeast ORFs when expressed as a fusion to green fluorescent protein at the C terminus. The authors speculate that this high rate of mislocalization may be the result of steric hindrance or interruption of critical C-terminal localization/retention sequences (30). Thus, if a portion of the 40% of mislocalized proteins was the result of steric hindrance, it is reasonable to assume that this steric hindrance would also affect protein-protein interactions. Finally by generating the allele libraries as entry clones, they may be transferred to any destination vector, allowing them to be used in any assay; the only requirement is the Gateway adaptation of the expression vector used to express the library. Generating allele libraries via gap repair does not allow for this flexibility.

Additional applications for the libraries include reverse two-hybrid analysis of DNA- and RNA-protein interactions whereby the nucleotide binding interfaces may be identified. Also, the libraries may be used to analyze protein interactions in mammalian systems, such as mammalian protein-protein interaction trap (MAPPIT) (31). Plus the libraries may be used in assays that select for enhanced, rather than disrupted, interactions. In contrast, generating allele libraries via gap repair limits the use of the library to yeast assays, and it can only be used a single time. This fact, in addition to the low library complexity and transformation efficiencies typical of gap repair, clearly illustrates the advantages of using the pDONR-Express system.

It should be pointed out that roughly half of the 5-FOA-resistant colonies recovered from our reverse two-hybrid screen had a wild type prey sequence. There are numerous mutations that can occur spontaneously that result in a URA^- phenotype, including mutations in the URA3 reporter gene itself, mutations in the Gal4p binding sites upstream of URA3, or mutations within the promoters or GAL4 sequences on the bait and prey plasmids themselves. Because there are more sites of potential inactivation of URA3 gene expression in a two-hybrid context, the frequency of spontaneous 5-FOA resistance is higher in reverse two-hybrid screens than for wild type S. cerevisiae strains. Because we have eliminated the major cause of false positive 5-FOA resistance, this spontaneous mutational frequency, formerly a minor cause of false positives, is now the major remaining cause of false positives.

From a systems approach, an analysis of networks is best conducted when individual edges (e.g. interactions) are knocked out rather than the node (e.g. protein). This is especially true for hub proteins, which interact with multiple partners. IDAs have been isolated that have lost the ability to interact with one partner but have retained interaction with others (12, 32, 33). With the recent publication of several metazoan protein interaction networks (5, 6, 8, 9), the introduction of IDAs into these cells/organisms and simultaneous knockdown of the wild type gene with techniques such as RNA interference have the potential to analyze protein interaction networks as never before. The ability to disrupt specific edges while maintaining nodes is only possible using IDAs, and an efficient approach to reverse two-hybrid screening has been a limiting factor in IDA isolation. Thus, our new strategy for conducting reverse two-hybrid screens should facilitate the rapid isolation of IDAs for use in protein-protein interaction network analysis.

In summary, we have developed a new method for allele library generation for reverse two-hybrid analysis of protein interactions that significantly reduces background (by 100-fold) typically associated with reverse two-hybrid analysis and expedites the isolation of IDAs, which allow the identification of interaction domains and interfaces. The effectiveness of the system was recently demonstrated by Kritikou et al. (34) who identified the interaction interface of MPK-1 that mediates interaction with GLA-3. Moreover this new method for allele library generation is applicable to any mutant library generation approach where truncated proteins are undesirable.

Acknowledgments—We thank Jon Chesnut, Louis Leong, James Fan, Antje Taliana, Isolde Kusser, and Beverly Schwabe for supplying plasmids; Katerina Kourentzi for the PureLink plasmid isolation protocol; Suzanne McNeely and Venna Parkhi for performing sequencing; David Mandelman for assistance with ViewerLite; and Paul Predki for reading the manuscript.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

† To whom correspondence should be addressed: Ambyr Genetics, 100 Columbia #200, Aliso Viejo, CA 92656. E-mail: phillipngray@gmail.com.

REFERENCES

1. Fields, S., and Song, O. (1989) A novel genetic system to detect protein-protein interactions. Nature 340, 245–246

2. Rain, J. C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachtner, V., Chernama, Y., Labigne, A., and Legrain, P. (2001) The protein-protein interaction map of Helicobacter pylori. Nature 409, 211–215

3. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Sinrivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Roth, J. M. (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature 403, 623–627

4. Itò, T., Chiba, T., Ozawa, R., Yoshihara, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactionome. Proc. Natl. Acad. Sci. U. S. A. 98, 4569–4574

5. Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chennubhotla, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. F., Lamesh, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. Y., Berriz, G. F., Jacotot, L., Vaglio, P., Reubel, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequeria, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., Van Den Heuvel, S., Piano, F., Vandenbunder, J., Sarstedt, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E., and Vidal, M. (2004) A map of the interactome network of the metazoan C. elegans. Science 303, 540–543

6. Giotta, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machinemith, H., Welsh, M., Kong, Y., Yershen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Iino, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carolla, S., Bickelhaupt, E., Lazovatsky, O., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., Jr., White, K. P., Braverman, M., Jarvis, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chait, J., and Rothberg, J. M. (2003) A protein interaction map of Drosophila melanogaster. Science 302, 1727–1736

7. Stanyon, C. A., Liu, G., Mangiola, B. A., Patel, N., Giot, L., Kuang, B., Zhang, H., Zhong, J., and Finley, R. L., Jr. (2003) Detection and modulation in vivo of helix-loop-helix protein-protein interactions. J. Biol. Chem. 288, 5–8

8. Gratas, A., Venkataraman, K., gene-sized DNA fragments using polymerase chain reaction. Anal. Biochem. 244, 337–353

9. Matsumura, I., and Ellington, A. D. (2002) Mutagenic polymerase chain reaction of protein-coding genes for in vitro evolution. Methods Mol. Biol. 182, 259–267

10. Lo Conte, L., Chothia, C., and Janin, J. (1999) The atomic structure of protein-protein recognition sites. J. Mol. Biol. 285, 2177–2198

11. Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., and O’Shea, E. K. (2003) Global analysis of protein localization in budding yeast. Nature 425, 686–691

12. Eckeyserman, S., Verhee, A., de Heyden, J. V., Lemmens, I., Ostade, V. X., Vandenkerchove, J., and Tavernier, J. (2001) Design and application of a cytokine-receptor-based interaction trap. Nat. Cell Biol. 3, 1114–1119

13. Itoyama, K., Shimizu, K., Chiba, N., Inoue, T., Maekawa, S., and Tomizawa, K. (1996) Mutational analysis of the yeast Saccharomyces cerevisiae: application of a differential interaction trap assay for examining protein-protein interactions. Genetics 147, 479–492

14. Jiang, R., and Carlson, M. (1996) Glucose regulates protein interactions within the yeast SNF1 protein kinase complex. Genes Dev. 10, 3105–3115

15. Vidal, M., and Harlow, E. (1999) Prospects for drug screening using the reverse two-hybrid system. Trends Biotechnol. 17, 374–381

16. Vidal, M., Braun, P., Chen, E., Boeke, J. D., and Harlow, E. (1996) Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. Proc. Natl. Acad. Sci. U. S. A. 93, 10321–10326