Avoiding the Ends: Internal Epitope Tagging of Proteins Using Transposon Tn7

Rebecca E. Zordan,*,[1] Brian J. Beliveau,* Jonathan A. Trow,* Nancy L. Craig,[†] and Brendan P. Cormack*,[2]
*Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2185, and [†]Department of Molecular Biology and Genetics, The Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2185

ABSTRACT Peptide tags fused to proteins are used in a variety of applications, including as affinity tags for purification, epitope tags for immunodetection, or fluorescent protein tags for visualization. However, the peptide tags can disrupt the target protein function. When function is disrupted by fusing a peptide to either the N or C terminus of the protein of interest, identifying alternative ways to create functional tagged fusion proteins can be difficult. Here, we describe a method to introduce protein tags internal to the coding sequence of a target protein. The method employs *in vitro* Tn7-transposon mutagenesis of plasmids for random introduction of the tag, followed by subsequent Gateway cloning steps to isolate alleles with mutations in the coding sequence of the target gene. The Tn7-epitope cassette is designed such that essentially all of the transposon is removed through restriction enzyme digestion, leaving only the protein tag at diverse sites internal to the ORF. We describe the use of this system to generate a panel of internally epitope-tagged versions of the *Saccharomyces cerevisiae* GPI-linked membrane protein Dcw1 and the *Candida glabrata* transcriptional regulator Sir3. This internal protein tagging system is, in principle, adaptable to tag proteins in any organism for which Gateway-adapted expression vectors exist.

KEYWORDS *Saccharomyces*; epitope tag; protein tagging; transposition

RECOMBINANT fusion proteins are essential tools in molecular studies across all organisms. Fusing fluorescent proteins to a protein of interest (POI) allows for direct visualization *in situ*. Fusing peptide epitopes, for which antibodies have been developed, to the POI is fundamental to many techniques, including Western blots, immunofluorescence, co-immunoprecipitation, chromatin immunoprecipitation, and purification (Arnau *et al.* 2006; Young *et al.* 2012; Bell *et al.* 2013).

We describe an approach to add a protein or peptide epitope to a POI. One challenge to epitope tagging is choosing the location to attach the epitope to the POI. The standard design appends the epitope to the N- or C-terminus of the POI. However, in some instances, proteins cannot be tagged at the N or C terminus because a tag interferes with function, disrupting, for example, trafficking or post-translational modification. Tags can also interfere with protein folding or structure and disrupt protein–protein interactions.

Our efforts to study the localization and function of Dcw1 in *Saccharomyces cerevisiae* have been hampered because we cannot introduce a tag at either terminus. Dcw1 and its paralog Dfg5 are important in cell-wall structure and integrity in *S. cerevisiae* and other fungi (Kitagaki *et al.* 2002, 2004; Gonzalez *et al.* 2010; Maddi *et al.* 2012). Traditional methods of N- or C-terminally tagging Dcw1 are expected to fail due to interference with protein localization because Dcw1 localization at the cell membrane requires an N-terminal secretory signal sequence and a C-terminal signal sequence for addition of a glycosylphosphatidylinositol (GPI) anchor. Modeled after published constructs (Kitagaki *et al.* 2002), we initially engineered an HA-tag located at residue 26 in *DCW1* (Dcw1-HA_{26}), just downstream of the N-terminal signal sequence, but found that this tagged protein, while viable, is only partially functional, conferring a temperature-sensitive phenotype. This motivated

doi: 10.1534/genetics.114.169482
Manuscript received August 8, 2014; accepted for publication February 23, 2015; published Early Online March 5, 2015.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.169482/-/DC1.
Sequence data from this article have been deposited with the GenBank Data Library under accession no. KP698385 (pRZ49 = Tn7-mEOS2); no. KP698386 (pRZ98 = Tn7-biotin); no. KP698387 (pRZ99 = Tn7-6His); no. KJ939358 (pRZ101 = Tn7-FLAG); no. KP698388 (pRA102 = Tn7-HA); no. KP698389 (pRZ103 = Tn7-myc); and no. KP698390 (pRZ106 = Tn7-GFP).
[1]Present address: Department of Biology, Notre Dame of Maryland University, Baltimore, MD 21210.
[2]Corresponding author: Johns Hopkins University School of Medicine, PCTB 522, 725 N. Wolfe St., Baltimore, MD 21205-2185. E-mail: bcormack@jhmi.edu

the development of a transposon-based system to introduce epitope tags throughout *DCW1*; the resulting library of epitope (3× FLAG)-tagged alleles of *DCW1* was then screened for function, permitting the isolation of functional internally tagged alleles of *DCW1*. To demonstrate the generality of the method, we mutagenized a second gene, *SIR3* from *Candida glabrata*. *SIR3* is not an essential gene but is absolutely required for subtelomeric transcriptional silencing in *C. glabrata* (De Las Penas *et al.* 2003). We used the transposon-based method to internally tag *CgSIR3* with four different protein tags, including the fluorescent proteins GFP and mEOS2.

Transposons have been used as tools to introduce epitope tags and fluorescent proteins randomly into targets (Merkulov and Boeke 1998; Ross-Macdonald *et al.* 1999; Manoil and Traxler 2000; Kumar *et al.* 2002; Sheridan *et al.* 2002; Kumar *et al.* 2004; Osawa and Erickson 2005). In some cases, the transposons are used to identify regions of the target POI into which a tag may be inserted without disrupting target function, requiring later cloning steps to insert the epitope tag at permissive sites (Spreghini *et al.* 2003). In a recently described system, TAGIT, Tn5 transposition is used to introduce cassettes containing epitope tags internally into target genes (Gregory *et al.* 2010). Resulting insertions are screened to identify in-frame fusions in the target gene. Excision of the bulk of the transposon is done *in vivo*, using Cre recombinase.

We describe a method, related to TAGIT, using transposon Tn7 and *in vitro* mutagenesis to introduce epitopes into two different open reading frames (ORFs), generating functional internally tagged alleles. Tn7 mini transposons have been shown to have very little sequence bias, making them ideal tools for random mutagenesis (Biery *et al.* 2000; Seringhaus *et al.* 2006; Green *et al.* 2012). Our plasmid-based epitope-tagging system generates large libraries of internally tagged ORFs, which can be screened for function to identify useful fusion proteins.

Materials and Methods

Plasmids used in this study are listed in Supporting Information, Table S1. DNA primers are listed in Table S2.

Media

Escherichia coli was routinely grown at 37° in LB media containing appropriate antibiotics for selection. For any media including trimethoprim, Oxoid Isosensitest media was used instead of LB. Carbenicillin, not ampicillin, was used to select for plasmids marked with the ampicillin-resistance gene. Antibiotics were added at the following final concentrations: carbenicillin (Car; 100 μg/ml), kanamycin (Kan; 30 μg/ml), and trimethoprim (Tmp; 10 μg/ml). Solid media for *E. coli* growth was supplemented with 1.5% agar.

S. cerevisiae and *C. glabrata* strains were typically grown at 30° on YPD media (10 g/liter yeast extract, 20 g/liter peptone, 2% dextrose). All solid yeast media contained 2% agar. To maintain His- and Ura-marked plasmids, SD-His (1.7 g/liter yeast nitrogen base without amino acids or ammonium sulfate, 5 g/liter ammonium sulfate, 1.92 g/liter SC-His amino acid mixture, 2% dextrose) or SD-Ura media (1.7 g/liter yeast nitrogen base without amino acids or ammonium sulfate, 5 g/liter ammonium sulfate, 6 g/liter casamino acids, 2% dextrose) were used. To maintain plasmids with a nourseothricin (NAT) marker, YPD was supplemented with 50 μg/ml NAT in liquid media or 100 μg/ml in solid media. To select against Ura-marked plasmids, 5-FOA media (1.7 g/liter yeast nitrogen base without amino acids or ammonium sulfate, 5 g/liter ammonium sulfate, 6 g/liter casamino acids, 25 mg/liter uracil, 1 g/liter 5-FOA, 2% dextrose) was used.

Strains and transformation

All *S. cerevisiae* strains used in this study are listed in Table S3; all *C. glabrata* strains are listed in Table S4. Standard lithium acetate transformation protocols were used (Hill *et al.* 1991). DH10 *E. coli* cells were used for routine cloning. Strain BW23473 was used to maintain the Tn7 donor plasmids, which have a R6Kγ origin (Metcalf *et al.* 1996). DB3.1 (Life Technologies) was used to propagate Gateway destination vectors containing the *ccdB* cassette. Highly competent MegaX DH10B T1R cells (Life Technologies) were used to maintain large libraries of isolates throughout the mutagenesis procedure.

BY240 strain construction

DCW1 was deleted in a clean two-step loopout from the *dfg5*ΔG strain from the yeast knockout collection (Winzeler *et al.* 1999). Because *DCW1* and *DFG5* are synthetically lethal (Kitagaki *et al.* 2002), the *dfg5*ΔG yeast strain was transformed with a His-marked plasmid carrying a wild-type copy of *DCW1*, prior to deletion of *DCW1*. The intergenic regions immediately flanking *DCW1* were amplified from genomic DNA using primers 1975–1978. The resulting fragments were cloned into YIPlac211 (*URA3*-marked). YIPlac211-*DCW1* was linearized with a *Kpn*I digest (separating the 5′ and 3′ flanking regions) and integrated at the *DCW1* genomic locus. YIPlac211-*DCW1* was replaced with a clean deletion construct, created by amplifying the flanking regions from YIPlac211 with primers 1975 and 1978. Clean deletions were identified by counterselection against the *URA3* marker in YIPlac211-*DCW1*. This completely removes the *DCW1* ORF from the genomic locus and leaves a *Kpn*I scar in its place. A plasmid shuffle replaces the *DCW1*(His) plasmid with pCU-*DCW1*, creating BY240.

Tn7-FLAG donor plasmid construction

The transposon FLAG donor plasmid is pRZ101, which carries Tn7-FLAG and is based on the suicide plasmid backbone pJP5603 (Penfold and Pemberton 1992), which contains the R6Kγ origin of replication (ORI). We modified pJP5603 by removing the *Xba*I site from the polylinker by treatment with Klenow and religation. Overall, the transposon was assembled modularly in other vector backbones and then subcloned into the modified pJP5603 backbone. Tn7L was amplified as a 236-bp fragment from pIC6 (Castano *et al.* 2003) using primers 3201 and 3202. This PCR product introduces an *Fse*I site at the distal

end of Tn7L and was cloned as a *Bam*HI-*Asc*I fragment. The FLAG epitope is a 3× FLAG tag flanked by flexible linkers. It was synthesized by DNA2.0 and is flanked by *Asc*I and *Xba*I restriction sites, allowing the FLAG tag to be easily subcloned into the Tn7 donor backbone. The *dhfr* gene was amplified from pAT-2 (Devine and Boeke 1994) cloned as an *Xba*I-*Pst*I fragment. The Tn7R end is derived from $Tn7R^{1\text{-}70*}$ in pMCB64 (Biery *et al.* 2000) (which includes a *Pme*I site at the distal Tn7R end) and was synthesized as a *Pst*I-*Eco*RI fragment (DNA2.0).

Other versions of the Tn7-tag donor vector were also constructed.The HA, biotinylation target sequence (bio), 6× His, and 3× myc epitope tags were all synthesized by DNA2.0. The mEOS2 epitope tag was PCR-amplified from pET28-ftsZ-mEOS (a gift of Jie Xiao) using primers 4724 and 4725. The GFP tag was PCR-amplified from pGRB2.3 (Zordan *et al.* 2013) using oligos 5300 and 5301 and then subcloned into the DNA2.0 backbone to position the GFP tag between *Asc*I and *Xba*I sites in the backbone. All tags were subcloned into the Tn7-tag donor backbone (as described for pRZ101) with the *Asc*I and *Xba*I sites flanking the epitope tag.

The following sequences for the Tn7-tag donor vectors are available from GenBank: pRZ49 = Tn7-mEOS2 (accession no. KP698385), pRZ98 = Tn7-biotin (accession no. KP698386), pRZ99 = Tn7-6His (accession no. KP698387), pRZ101 = Tn7-FLAG (accession no. KJ939358), pRA102 = Tn7-HA (accession no. KP698388), pRZ103 = Tn7-myc (accession no. KP698389), and pRZ106 = Tn7-GFP (accession no. KP698390).

DCW1 plasmid construction

pCU-DCW1: This plasmid is used in BY240 to cover the synthetic lethality between *DCW1* and *DFG5* gene deletions. It is derived from p416GPD (Mumberg *et al.* 1995), containing the *TDH3* (GPD) promoter, *CYC1* transcription terminator, and CEN/ARS and *URA3* markers for maintenance and selection in *S. cerevisiae*. The *DCW1* ORF was PCR-amplified from *S. cerevisiae* genomic DNA using primers 1629 and 1630 and subcloned into p416GPD using *Bam*HI and *Eco*RI.

DCW1-DONR vector (target): The *DCW1* gene was amplified from genomic DNA using oligos that appended standard attB1 (5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTA-3′) and attB2 (5′-GGGGACCACTTTGTACAAGAAAGCTGGGTA-3′) sites onto the forward and reverse oligos, respectively. The product was recombined into a Gateway entry vector pDONR201 using a standard Gateway BP clonase II reaction. The sequence of the plasmid was confirmed.

DCW1-destination vector: The entire intergenic region upstream of the *DCW1* ORF was PCR-amplified using primers 6333 and 6334 and cloned into a *ccdB*-containing backbone using *Sac*I and *Xba*I restriction sites present in the primers. A *Sac*I-*Xho*I digest was used to subclone the promoter and *ccdB* region into the p413GPD backbone (Mumberg *et al.* 1995), which contains the *CYC1* transcriptional terminator and a *HIS1* auxotrophic marker for selection in yeast.

Wild-type DCW1 expression vector: Gateway LR clonase II (Invitrogen protocols) was used to recombine the *DCW1*-DONR vector and *DCW1*-destination vector, creating a wild-type, untagged version of the *DCW1* expression vector pRZ160.

Dcw1-HA$_{26}$ vector: This vector is a derivative of pRZ160 with an HA tag inserted at amino acid position 26.

Verify DCW1 expression vector functions

Before beginning mutagenesis, it was important to test whether the destination vector and target gene would function in our system. We sequence-verified the wild-type *DCW1* expression vector pRZ160 and then transformed it into BY240 and selected for growth on SC-His media (selecting for pRZ160). The resulting strain was grown on 5-FOA to verify that the strain could lose the original pCU-*DCW1* plasmid and survive with the new *DCW1*(His) plasmid, pRZ160. This confirms that the *DCW1* promoter and *DCW1* ORF functioned in *S. cerevisiae* and would be suitable substrates for mutagenesis.

Transposition reaction and processing pool

A detailed transposition and processing protocol is in File S1. In short, TnsA, TnsB, and $TnsC^{A225V}$ enzymes (purified as described in Gamas and Craig 1992 and Choi *et al.* 2013) were used in an *in vitro* reaction, mobilizing the Tn7-tag cassette from a Tn7 donor vector into a target vector containing the *DCW1* ORF or the *CgSIR3* ORF. The mutagenized plasmids were recovered by transforming into *E. coli* MegaX cells and selecting for appropriate drug resistance. The *DCW1* ORF was mobilized using Gateway recombination enzymes in a dedicated *DCW1* destination vector, creating a mutagenized *DCW1* expression pool; likewise, following mutagenesis, the *CgSIR3* ORF was mobilized into a corresponding *SIR3* destination vector. These expression pools were recovered in *E. coli* MegaX cells and selected in sequential rounds of Car and Tmp selection. The left end of Tn7 was removed by *Fse*I digestion; the right end of Tn7 and the Tmp^R marker were removed by *Pme*I digestion.

Screening and analysis of Dcw1-FLAG alleles in S. cerevisiae

The final *DCW1*-FLAG plasmid pools were transformed into *S. cerevisiae* and grown on SD-His plates at 30° for 2 days. Transformants on the SD-His plates were replica-plated to 5-FOA plates and grown at 37° for 1 day and then re-replica-plated to 5-FOA plates and grown for 1 day at 37°. Only cells with functional *DCW1*-FLAG alleles will grow; those with nonfunctional *DCW1*-FLAG alleles will die on 5-FOA because of counterselection against the pCU-*DCW1* plasmid.

We performed colony PCR on functional FOA^R transformants to qualitatively determine where the FLAG was inserted within the *DCW1* ORF. The PCR to determine FLAG insertion position used primers 2766 and 6244 or primers 5160 and M13F. We also performed PCR with 2766 and a primer (5626) that reads out of Tn7L; any isolates that gave a PCR product were eliminated from further study. A subset of isolates predicted to have different sites of insertion (based on PCR product size) were

sequenced with primer 6244 to identify the exact placement of the FLAG tag. We screened 276 isolates by colony PCR, sent 48 isolates for sequencing, and found 10 unique DCW-FLAG alleles with internal insertions.

For the 10 unique isolates, functional *DCW1*-FLAG plasmids were isolated from *S. cerevisiae* and recovered by transformation into *E. coli* (Hoffman 2001). The plasmids were individually retransformed into BY240, and their growth phenotype was confirmed. These 10 clean functional strains were used for subsequent growth analysis and for Western blots. The 10 clean *DCW1*-FLAG strains, the *DCW1*-HA strain, and control strains were streaked onto YPD, and complementation was tested again on 5-FOA (at 39°, 37°, and 30°), SC-Ura, SC-His, and YPD plates at 30°. After growth for 2 days, pictures were taken on an Alpha Imager and growth was compared (Figure 4).

Construction of C. glabrata SIR3 Gateway vectors

The *C. glabrata SIR3* destination vector was constructed by PCR-amplifying the regions flanking the *SIR3* ORF from strain CBS138 (Dujon *et al.* 2004). Primers 3220 and 3221 were used to amplify the 5′ flanking region; primers 3222 and 3223 were used to amplify the 3′ flanking region. These PCR products were subcloned into a Gateway backbone on either side of a *ccdB* cassette, using the restriction sites present in the primers. This Gateway backbone contains replication origins and and an Amp^R cassette for selection in *E. coli*, as well as a CEN/ARS and Nat^R marker for maintenance and selection in *C. glabrata*. The *SIR3* entry vector was created by PCR-amplifying the SIR3 ORF from *C. glabrata* strain BG2 (Cormack and Falkow 1999) using primers 4470 and 4484 and subsequently introducing the ORF into the Gateway backbone pDONR201 using a Gateway BP reaction (Life Technologies). We note that there are three polymorphisms in the *SIR3* ORF, relative to the CBS138 sequence, but these do not affect function, as a wild-type *SIR3* expression vector generated from this ORF allele (pRZ47) complements a *sir3Δ* defect in *C. glabrata*. Schematic drawings of the SIR3 entry vector and destination vector are shown in Figure S2.

Microscopy of Sir3-GFP strains

Eight *C. glabrata* strains carrying various Sir3-GFP alleles, as well as a negative control strain carrying an untagged Sir3 vector, were grown to stationary phase in liquid YPD+Nat. Cells were washed in PBS and resuspended in PBS, and 5 μl of the cells was mounted on a slide. Images were taken using a Zeiss Axioskop microscope with a 100× objective. The captured images are automatically displayed with optimized brightness and contrast settings in Image J; these maximum and minimum values for the entire captured image are listed in Figure 8. Image J software was used to adjust the constrast to the same settings for all images, thereby allowing more direct comparison of GFP brightness across all images. The adjusted contrast images were converted to 8-bit images, cropped, and resized using Adobe Photoshop and Adobe Illustrator.

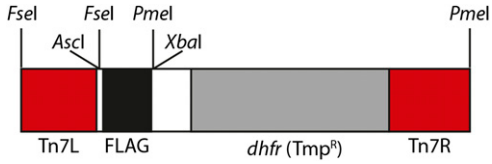

Figure 1 Schematic of the Tn7-FLAG construct. Areas indicated by red are Tn7 left and right ends. The 3×FLAG epitope tag is indicated by black. The gray area is the *dhfr* gene, which confers resistance to trimethoprim (Tmp^R). Restriction sites of interest are indicated above the schematic.

S. cerevisiae protein preparation and Westerns

S. cerevisiae strains carrying various *DCW1*-FLAG plasmids, or a wild-type *DCW1* untagged plasmid (negative control), were grown in SD-His media to mid-log phase (OD_{600} between 0.1325 and 0.2325). Cells were pelleted by centrifugation and stored at −80°. Cell pellets were weighed on a microscale; these weights were used later to normalize loading between strains. Lysates were prepared similar to the method of Frieman and Cormack (2004).To prepare lysates, resuspend cell pellets in 50 mM Tris (pH 8.0) supplemented with protease inhibitors (Roche, 04693132001) and lyse with glass beads using a FastPrep (Bio101 Thermo Savant) (three times: 45 sec, 6.5 speed, ice 1 min between beatings). Supernatants were transferred to clean tubes, and beads were washed in 400 μl Tris (pH 8.0) + protease inhibitors and combined with an earlier fraction (800 μl lysate total). Lysates were clarified by centrifugation at 13,000 × *g*, 10 min, 4° in a microfuge. Supernatant was saved as the "cytoplasmic" fraction. The pellet was resuspended in 1 ml 50 mM Tris (pH 8.0) + 2% SDS and boiled for 20 min, vortexing to mix every 10 min. After centrifugation (13,000 × *g*, 10 min, 4°), the supernatant was saved as the "plasma membrane" fraction. The remaining pellet material was resuspended, washed in Tris+SDS, boiled, and centrifuged (as above) three additional times; supernatants were discarded. The pellet was washed four times (twice in 1 ml, twice in 500 μl) in 50 mM Tris + protease inhibitors and spun as before. The pellet was washed once in 500 μl 33 mM potassium phosphate and then resuspended in 100 μl 33 mM potassium phosphate + 60 mM β-mercaptoethanol; this is the "cell-wall" fraction.

Material from each fraction was loaded onto an SDS-PAGE gel; volumes were normalized across strains using pellet weights. Proteins were transferred to Immobilon-P membrane and immunoblotted using antibodies as indicated. Mouse monoclonal α-FLAG (F1804, Sigma) was used at 1:2000 diluted in TBS+3% milk; α-Pgk1 (A6457, Molecular Probes) was used at 1:1000 in TBS+3% milk. Both of these were used in conjunction with α-mouse HRP-linked (Cell Signaling) secondary antibody at a 1:2000 dilution. The α-Dcw1 polyclonal antibody was raised against the peptide VELDLDNYESLQ, representing amino acids 22–33 in the Dcw1 protein (Covance, Princeton, NJ). For Dcw1 detection, the α-Dcw1 primary antibody was diluted at 1:1000 in TBS+3% milk, and the secondary antibody was a 1:5000 dilution of α-rabbit HRP-linked antibody (Cell Signaling). Amersham's ECL kit (RPN2132) was

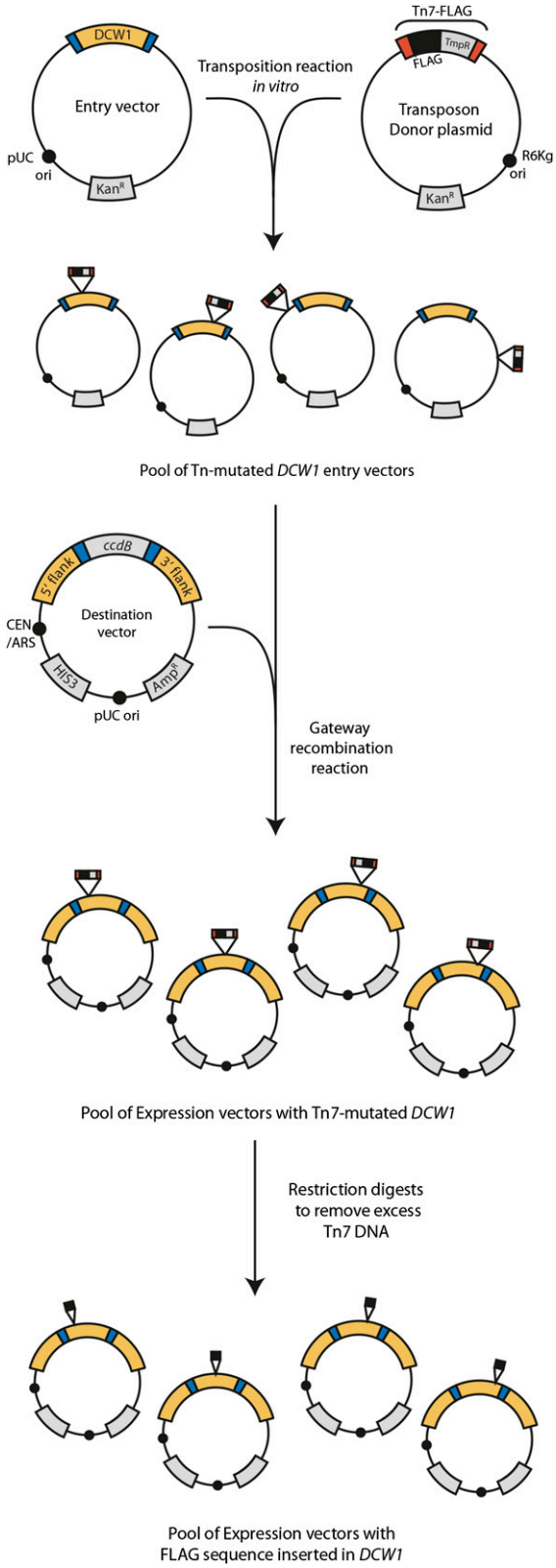

Figure 2 Overview of mutagenesis work flow. An *in vitro* mutagenesis reaction introduces the Tn7-FLAG transposon randomly throughout the target entry vector (here, *DCW1*). Mutagenized *DCW1* plasmids are selected on the basis of Tmp^R and Kan^R. To isolate only those versions that have a Tn7-FLAG inserted in *DCW1*, and not in the backbone, a Gateway LR reaction is used to mobilize the *DCW1* gene into a *DCW1* destination vector. After transformation into MegaX *E. coli* cells and selection with Tmp and Car, only expression vectors that have a mutagenized *DCW1* gene should remain. Restriction digests are used to remove the bulk of the transposon: *Fse*I removes the Tn7L end; *Pme*I removes the Tmp^R cassette and Tn7R. After this series of digestions, the pool will consist of *DCW1* expression vectors with a 3× FLAG epitope inserted randomly, in either direction, in the *DCW1* gene.

used for chemiluminescent detection; the α-FLAG and α-Pgk1 blots were exposed to film for 5–10 min; the α-Dcw1 blots were exposed for 45 min.

C. glabrata protein preparation and Westerns

C. glabrata strains carrying different Sir3-GFP plasmids or a wild-type *SIR3* untagged plasmid (negative control) were grown in YPD+Nat media to log phase (OD_{600} between 0.5 and 0.635). Cells were pelleted by centrifugation and stored at −80°. Lysates were prepared in a urea lysis buffer (Ubersax *et al.* 2003), supplemented with protease inhibitors (Roche, 11836170001) . Proteins were separated on a 3–8% Tris–acetate SDS-PAGE gel (Nupage) and transferred to an Immobilon-P PVDF membrane. A polyclonal α-GFP primary antibody (Abcam, ab290) was used at 1:5000 dilution in PBST+5% milk; the α-rabbit HRP-linked secondary antibody (GE Healthcare, NA934V) was used at a 1:5000 dilution. Chemiluminescent detection was performed using the Amersham ECL kit (RPN2132).

Viability selection and sequence analysis of functional tagged SIR3 alleles in C. glabrata

SIR3 was mutagenized four separate times, using the Tn7-bio, Tn7-myc, Tn7-mEOS2, or Tn7-GFP donor vectors. After processing the mutagenized plasmid pools to remove leftover Tn7 sequence, the final Sir3-bio, Sir3-myc, Sir3-mEOS2, and Sir3-GFP pools were transformed into *C. glabrata* strain CGM293, and transformants were selected by growth on YPD+Nat plates for 2 days at 30°. Strain CGM293 carries *URA3* integrated into a subtelomeric location, where it is subjected to Sir3-dependent transcriptional silencing. Only cells that carry a functional tagged version of Sir3 have intact subtelomeric silencing, thus silencing the URA3 gene at the telomere and permitting growth on plates containing 5-FOA. To select functional clones of Sir3, the YPD+Nat plates were replica-plated onto 5-FOA plates and grown for 2 days at 30°. For the Sir3-myc transformation the pool was outgrown in liquid YPD+Nat prior to plating. This pool showed skewed representation of particular insertion sites. Subsequently, for the other three pools, the transformations were plated directly onto YPD+Nat plates, and these pools show a more even representation of different functional alleles. As with the *DCW1* screening, we performed colony PCR to qualitatively determine where the tag had inserted into *SIR3*. Judging from these PCR product sizes, plasmids carrying SIR3 with a range of insertion sites were sent for sequencing to determine the exact position of the tag within *SIR3* (Table S6 and Table S7).

Table 1 Sizes of mutagenized *DCW1* pools

Pool	Processing step	No. of colonies (calculated to full reaction size)					Fold Coverage	Fold coverage based on which drugs?
		Tmp^R Kan^R	Tmp^R	Kan^R	Car^R	Tmp^R Car^R		
Mutagenesis	Mutagenesis of entry vector	2.01E+05	2.45E+05	5.00E+06	ND	ND	47.5×	Tmp Kan
xpC[a]	Gateway LR to create expression pool	1.20E+04	ND	8.50E+04	5.00E+06	2.5E+05[a]	91×	Tmp Car
xpT	xpC, retransformed and selected for Tmp	9.00E+03	4.00E+06	2.55E+05	1.05E+08	4.00E+06	~91×	Tmp Car
xpT-Tn7L	After removal of Tn7L	5.60E+03	3.06E+07	1.64E+04	3.10E+07	2.74E+07	~91×	Tmp Car
xpT-Tn7L-Tn7R	after removal of Tn7R	2.00E+02	1.87E+06	1.07E+05	5.40E+07	2.10E+06	~91×	Car

Colony counts, as plated after recovery during transformation, prior to drug selection in pool. Cells were diluted appropriately to prevent a lawn of growth and plated onto media as indicated. Numbers in the table are calculated to represent the number of colonies in the full transformation resistant to the given drug. ND, not determined.
[a] xpC colony counts were revised; see File S3 and Table S5. Fold coverage was calculated by comparing the number of colonies from the indicated drug selection to the size of the DNA available for mutagenesis (File S2).

Results

Transposon design

The Tn7-FLAG transposon (Figure 1) is based off the miniTn7 design (Biery *et al.* 2000), using truncated Tn7 left and right (Tn7L and Tn7R) ends. The Tn7L region is 206 bp long, whereas the Tn7R region is 71 bp long. Previous research has shown the Tn7 ends can accommodate *Pme*I sites and still mobilize during *in vitro* transposition reactions (Biery *et al.* 2000). Here, we engineered *Fse*I and *Pme*I sites at the distal Tn7 transposon ends to facilitate removal of the bulk of the transposon sequence using restriction digests after the mutagenesis occurs (see later description). These restriction enzymes have 8-bp restriction sites, which occur rarely in the DNA genome sequence. The Tn7-FLAG construct contains two *Fse*I sites, which flank the Tn7L end; two *Pme*I sites flank a trimethoprim resistance (Tmp^R) cassette and the Tn7R end. *Pm*FseI digestion, therefore, removes the left end of the transposon, while PmeI digestion removes the right portion of the transposon. The transposon contains a 3× FLAG epitope tag, located between the internal *Fse*I and *Pme*I sites. The entire Tn7-FLAG construct is carried on the plasmid pRZ101, which serves as a Tn7 donor plasmid. The pRZ101 backbone includes a kanamycin resistance marker (Kan^R) and replicates using a R6Kγ ORI. This ORI functions only in *E. coli* cells containing the Π protein, encoded by the *pir* gene (Kolter *et al.* 1978; Haldimann *et al.* 1996; Metcalf *et al.* 1996).

During Tn7 mutagenesis, the Tn7 transposase, containing TnsA and TnsB subunits, creates double-strand breaks at the Tn7 ends, releasing the Tn7-FLAG cassette from pRZ101. $TnsC^{A225V}$ mediates insertion of the Tn7-FLAG cassette into a variety of targets with essentially no sequence specificity (Green *et al.* 2012). During host-mediated repair of the insertion product, a 5-bp duplication is introduced at the insertion site. The epitope tag is positioned within Tn7-FLAG so that it will be translated correctly in only one reading frame after subcloning to excise transposon sequences. All other reading frames contain STOP codons. Thus, five of six transposon mutants are expected to be nonfunctional as a consequence of the epitope tag being out of frame and generating an internal stop codon. After all mutagenesis and processing of the transposon mutants, the Tn7-FLAG cassette introduces 159 bp into the target (including the 5-bp duplication), resulting in a 53-amino-acid insertion internally fused into the target protein.

Tn7-FLAG can be easily adapted to introduce other epitope tags. Unique *Asc*I and *Xba*I sites flank the FLAG cassette, and these sites can be used to change the epitope tag to any other tag of interest. In our own work, we generated Tn7-6His, Tn7-myc, Tn7-GFP, Tn7-mEOS2, Tn7-biotin, and Tn7-3×HA donor plasmids, all of which have been confirmed to mobilize during *in vitro* transposition reactions (not shown).

Methodology overview

The *in vitro* mutagenesis method to introduce internal epitope tags is outlined schematically in Figure 2. To illustrate the utility of this system, we chose to internally tag the protein Dcw1, a GPI-anchored protein, which cannot be functionally tagged at either the N- or C-terminus. The method consists of three separable steps, all of which are carried out *in vitro*, with recovery of reaction products by transformation into *E. coli*. First, transposition is used to generate a pool of Tn7 insertions throughout the plasmid carrying the gene to be tagged. Second, Tn7 insertions within the ORF are isolated by a Gateway-mediated recombination step. Third, this pool of Tn7-mutated ORFs is treated with restriction enzymes to remove essentially all transposon sequences, leaving just the introduced protein tag.

Transposition: Using a target plasmid containing the *DCW1* ORF cloned into a Gateway entry vector, we performed an *in vitro* mutagenesis reaction with recombinant Tn7 transposase proteins. The reaction mobilizes the Tn7-FLAG cassette from the Tn7 donor plasmid into the *DCW1* target plasmid. This creates a pool of mutagenized *DCW1* entry vectors. This pool is transformed into highly competent *E. coli* DH10 cells and selected for both Tmp^R and Kan^R in liquid culture. The use of DH10 cells selects against the Tn7-FLAG donor plasmid, as this plasmid's R6Kγ ORI will not replicate in DH10 cells. Double drug selection selects against unmutagenized *DCW1* entry

vectors. Thus, after growth overnight in Tmp and Kan, only cells carrying mutagenized *DCW1* entry vectors should survive.

Isolation of mutagenized ORFs by Gateway recombination: The pool consists of plasmids that have a transposon inserted in the *DCW1* ORF as well as plasmids where the transposon is inserted in the plasmid backbone (and which have an unmutagenized ORF). To isolate only those ORFs that have been insertionally mutagenized, we used Gateway recombination to mobilize the *DCW1* ORF from the entry vector into a *DCW1* destination vector (Figure 2) to generate what we refer to as the expression pool. The *DCW1* destination vector contains *DCW1* promoter and terminator regions, flanking a *ccdB* cassette; it also contains replication origins and markers for propagation and selection in both *E. coli* and yeast. This *DCW1* expression pool was recovered by transformation into *E. coli* DH10 cells. Expression pool plasmids, with the transposon-containing *DCW1* ORF cassette, confer resistance to both Tmp and Car. We found that this double selection was maximally effective if we performed the drug selections sequentially (data not shown); accordingly, the transformed cells were initially selected only for carbenicillin resistance (Car^R), which selects for all destination vectors carrying a mutagenized or unmutagenized *DCW1* ORF. Entry vectors and destination vectors that did not recombine are selected against based on the Car^R and *ccdB* counterselectable cassette, respectively. After selecting for Car^R, plasmid DNA is recovered from the expression vector pool and transformed again into *E. coli* DH10 cells, this time selecting for Tmp^R. After growth in Tmp, only cells carrying *DCW1* expression vectors with a Tn7-FLAG cassette remain. Note that the Tn7-FLAG may be inserted in any reading frame at this point, as the pool of mutants has not been screened for function in yeast yet.

Removal of transposon sequences: Removal of the bulk of the transposon, leaving only the epitope tag, requires two restriction digests and ligation steps. These cloning steps are performed on the pooled DNA, allowing for easy processing of the entire insertion library. Digestion with *Fse*I releases the Tn7L end from each plasmid, which is then recircularized by ligation. This "–Tn7L" pool of plasmid DNA is transformed into *E. coli*, and the liquid culture is treated with Tmp and Kan simultaneously. This second round of Tmp selection is possible since the Tn7R end and *dhfr* gene remain in the construct and selects against any residual unmutagenized *DCW1* expression vectors that have escaped the previous selection step. Next, a *Pme*I digest is used to remove the *dhfr* (Tmp^R) gene and the Tn7R end (Figure 1). After intramolecular ligation, the plasmid pool is transformed into *E. coli* and selected for Car^R. The DNA isolated from this pool contains mutagenized *DCW1* expression vectors with the 59-aa FLAG epitope inserted in any of the six reading frames.

Monitoring pool sizes and complexity

Transposition efficiency was monitored by plating on various selective media. Transformation of our mutagenesis reaction resulted in 5×10^6 Kan^R colonies, representing all plasmids whether mutagenized or not. Of these, 2×10^5 (4%) contained transposon insertions (Tmp^RKan^R) (Table 1). This represents a 47-fold coverage of all possible insertion sites for the plasmid (File S2).

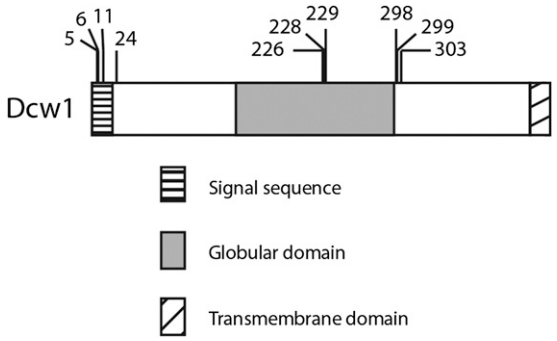

Figure 3 Location of FLAG insertions in functional *DCW1*-FLAG alleles. Domain structure of Dcw1 is based on predictions from ELM. The locations of FLAG insertions within the functional *DCW1*-FLAG alleles used in subsequent experiments are indicated by hash marks above the diagram. Dcw1 is 499 amino acid residues in length.

The efficiency of Gateway LR mobilization, as well as the fraction of ORFs carrying the Tn7-flag, was estimated by scoring for Car and Tmp resistance (File S3). We determined that the initial expression pool (xpC) contained 2.5×10^5 Tmp^RCar^R independent recombinants (Table 1). Thus, the majority of the pool complexity is maintained in the expression pool. For subsequent steps, *Fse*I digestion and *Pme*I digestion, we recovered at least 4×10^6 transformants, again ensuring that pool complexity was maintained through different cloning steps.

Complexity of the pool was qualitatively assessed using restriction digests of pool DNA at each step of the processing (File S4). Additionally, the plasmids from a subset of the final mutagenized expression pool were isolated and sequenced with primer 6244 to assess the distribution of Tn7 insertion sites in the plasmid. From 48 sequenced isolates, we found 29 had FLAG inserted in the reverse orientation (representing 18 unique sites) and 19 had FLAG inserted in the forward direction (representing 18 unique sites). For two nucleotide positions, we isolated FLAG insertions in both the "forward" and the "reverse" orientations in different isolates.

Selection for functional clones in yeast

The final *DCW1*-FLAG pool was transformed in batch into *S. cerevisiae* to screen for functional clones. Mutations in *DCW1* and the related gene *DFG5* are synthetically lethal in *S. cerevisiae* (Kitagaki *et al.* 2002). We use the strain BY240, which is a *dcw1*Δ *dfg5*Δ strain carrying a wild-type copy of *DCW1* on a *URA3*-marked plasmid (pCU-URA3). The plasmid backbone of the *DCW1*-FLAG expression pool contains a yeast CEN/ARS sequence and a *HIS3* marker for selection. After transforming the *DCW1*-FLAG pool into BY240, we selected against the original wild-type pCU-URA3 plasmid by growth in the presence of

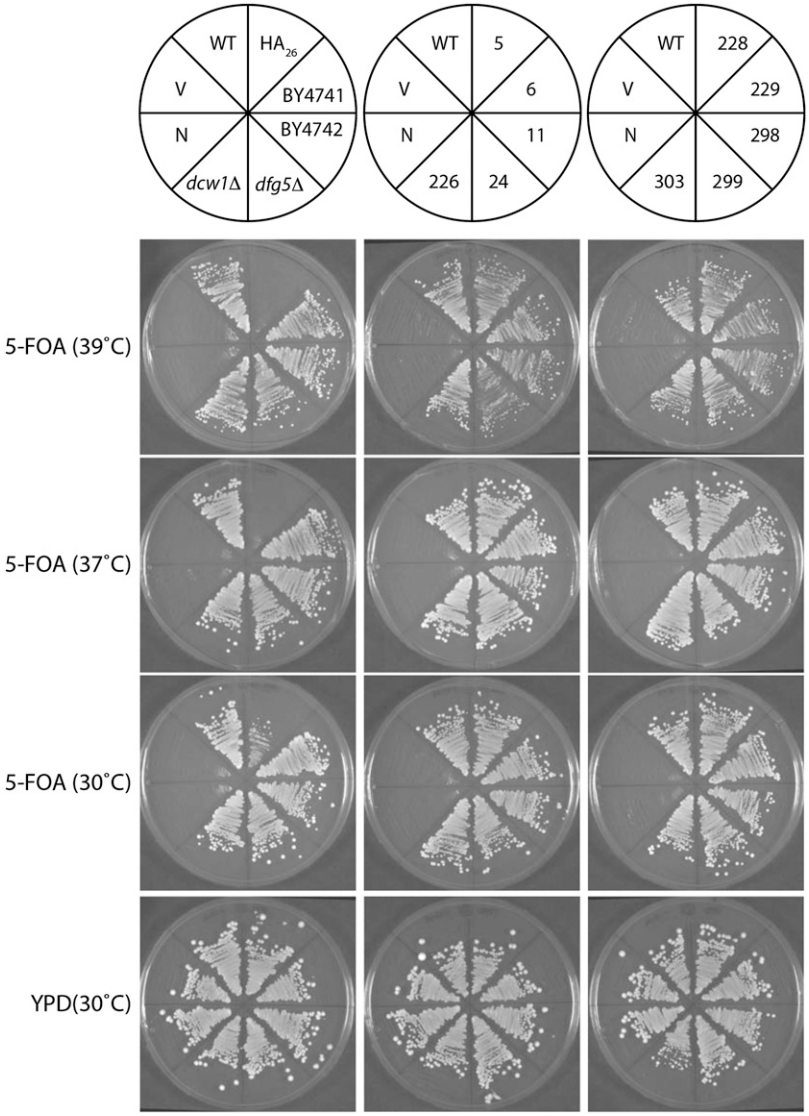

Figure 4 Select *DCW1*-FLAG isolates rescue synthetic lethality of *dcw1*Δ *dfg5*Δ mutants. Each yeast strain was grown on 5-FOA at three temperatures to test for Dcw1 function. Strains were also grown on YPD to verify overall viability. The positions of each strain are shown in the schematic at the top. Three control strains, labeled here as N, C, and WT, were streaked onto each experimental plate. Strain "N" is a negative control: the parental strain BY240, a *dcw1*Δ *dfg5*Δ pCU-*DCW1* strain derived from BY4742. Strain "V" is another negative control: BY240 transformed with an empty (His) vector, pRZ159. The "WT" strain is a positive control: strain BY240 transformed with a wild-type *DCW1* (His) plasmid. BY4741 and BY4742 are *DCW1 DFG1 ura3*Δ strains used to derive other strains. *dfg5*Δ and *dcw1*Δ are deletion mutants derived from BY4741 from the yeast knockout collection (Winzeler *et al.* 1999). The strain BY240 transformed with a *DCW1*-HA_{26} (His) plasmid is in the position labeled "HA_{26}." The remaining strains are BY240 transformed with various *DCW1*-FLAG (His) plasmids and labeled to indicate the amino acid position of the FLAG insertion.

5- 5-FOA. Only strains transformed with a functional *DCW1*-FLAG allele will grow on 5-FOA plates. We performed the 5-FOA selection at 37°; the elevated temperature is a mild cell-wall stress and ensures that any cells that grow have a fully functional *DCW1*-FLAG allele. We screened ~100,000 transformants, of which ~5% had functional versions of *DCW1*-FLAG (data not shown).

Transformants were first screened for presence of a FLAG insertion using whole-cell PCR (data not shown). Products had a range of sizes, representing FLAG insertions across the *DCW1* ORF. A subset of these was chosen for further characterization. Plasmids carrying functional *DCW1*-FLAG alleles were isolated from individual *S. cerevisiae* transformants and recovered in *E. coli*. The location of the FLAG insertion within the *DCW1* ORF was determined by sequencing. A number of unique insertion sites were identified, as illustrated on the schematic shown in Figure 3. The majority of insertions were, as expected, in the ORF. In addition, several isolates were identified in which the FLAG was inserted just upstream or downstream of the *DCW1* ORF in the sequence between the Gateway recombination sites and the ORF in the expression vectors (data not shown). We found that the Gateway attL1 and attL2 recognition sequences, required for mobilization of the *DCW1* ORF into the destination vector, are partially permissive such that certain Tn7-disrupted attL1 and attL2 sequences still, surprisingly, function in the Gateway reaction. These few insertions, therefore, have the epitope outside of the ORF sequence and represent contamination in the overall library.

Unique *DCW1*-FLAG plasmids were retransformed into *S. cerevisiae* strain BY240, and function was confirmed by testing for growth on 5-FOA plates (Figure 4). All internally tagged *DCW1*-FLAG isolates tested grew robustly on 5-FOA at temperatures up to 39°, confirming that the *DCW1*-FLAG alleles are fully functional. The *DCW1*-HA_{26} allele showed reduced growth on 5-FOA at 30° and no growth at 37° and 39°, indicating that it is a partially functional hypomorphic allele.

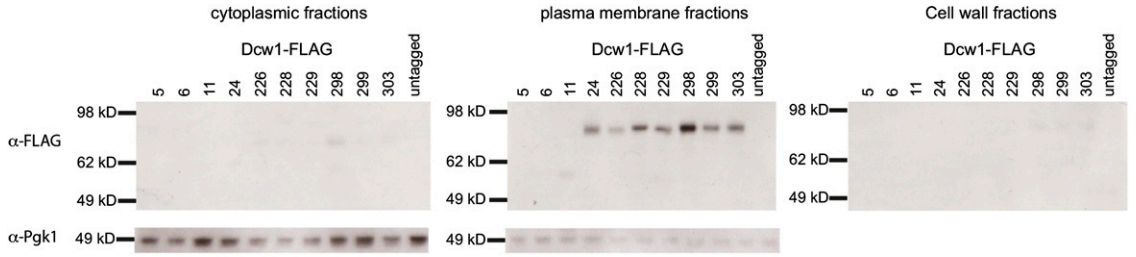

Figure 5 Western blot analysis of Dcw1-FLAG expression in *S. cerevisiae*. (Top) Anti-FLAG immunoblots of *S. cerevisiae* strains expressing Dcw1-FLAG alleles, representing the cytoplasmic, plasma membrane, and cell-wall fractions from each strain. Material from 7.5 × 10^5 cells is loaded for each strain in each fraction. Wells are labeled with the amino acid site of the FLAG tag within Dcw1. (Bottom) Anti-Pgk1 immunoblot of the cytoplasmic and plasma membrane fractions from strains carrying Dcw1-FLAG alleles. Cytoplasmic material from 3.75 × 10^5 cells is loaded for each strain. Membrane material from 3.5 × 10^5 cells was loaded for each strain

Detecting FLAG by Western blot

To verify that the FLAG epitope introduced into Dcw1 is detectable by Western analysis, we monitored Dcw1-FLAG protein levels in our strains carrying functional *DCW1*-FLAG clones. Cytoplasm, plasma membrane, and cell-wall fractions were isolated from log-phase cultures of each strain. Dcw1 is a membrane-bound GPI protein (Kitagaki *et al.* 2002), and abundant FLAG signal is easily detectable in the cell membrane fraction in 8 of the 10 strains expressing Dcw1-FLAG alleles, consistent with the fact that these are fully functional Dcw1-FLAG alleles (Figure 5). Untagged Dcw1, which would not be visible on the α-FLAG immunoblot, is predicted to have a molecular weight of 49.5 kDa. Dcw1-FLAG fusion proteins are expected to have a molecular weight of 55 kDa. However, mature Dcw1 is glycosylated, and previous studies have shown that Dcw1-HA has an apparent molecular weight of 80 kDa (Kitagaki *et al.* 2002), similar to what we observe in Figure 5. For three alleles, the FLAG insertion (at residue 5, 6, or 11) is predicted to disrupt the signal sequence (Figure 3). For two of these three, Dcw1-$FLAG_5$ and Dcw1-$FLAG_6$, the immunoblot showed no signal, and for one, Dcw1-$FLAG_{11}$, the protein was detectable in the membrane fraction, but its molecular weight is lower than expected (~60 kDa, instead of ~80 kDa).

To confirm that the localization observed of the Dcw1-FLAG fusion proteins represents true localization of untagged Dcw1, we also performed immunoblots using these same cellular fractions with an α-Dcw1 antibody raised against a peptide portion of native Dcw1 (Figure S1). The signal is much fainter than what we observed using α-FLAG antibodies; importantly, the Dcw1-FLAG fusion proteins show similar size distributions using both the anti-Dcw1 and anti-FLAG antibodies.

Mutagenesis of a second target gene, C. glabrata SIR3

The demonstrate the generality of the tagging methodology, we used the same approach to mutagenize the *SIR3* gene of *Candida glabrata*, which is required for silencing of sub-telomeric regions of the *C. glabrata* genome (De Las Penas *et al.* 2003). Schematic diagrams of the Sir3 Gateway constructs are shown in Figure S2. We used four different Tn7 derivatives (Tn7-myc, Tn7-GFP, Tn7-mEOS2, Tn7-biotin) to generate libraries of Tn7 insertions with between 3 and 6 × 10^5 independent insertions in each pool. Functional tagged alleles were identified by complementation of a *sir3Δ* strain. We sequenced the insertion sites for fully functional alleles, identifying tag insertion sites distributed across the ORF (Figure 6, Table S6, and Table S7). We chose to further characterize eight GFP-tagged *SIR3* alleles, which complement for the telomeric silencing defect (Figure S3), representing insertions sites distributed across the ORF. By Western analysis, all eight were expressed at approximately similar levels (Figure 7). For all eight alleles, Sir3-GFP was localized in puncta, as has been reported previously for wild-type Sir3 protein in *S. cerevisiae* (Cockell *et al.* 1995) (Figure 8).

Discussion

Here we describe a robust system for creating a library of internally tagged target protein constructs. This mutagenesis is carried out *in vitro*, and all subsequent steps of selecting for insertional mutants and removing excess Tn7 sequence are carried out on the bulk pool *in vitro*. This batch processing greatly reduces the effort in maintaining and screening the thousands of possible mutants.

As proof of principle, we mutagenized a target gene, *DCW1*, to saturation, introducing a 3× FLAG epitope throughout the ORF. Screening in yeast identified fully functional tagged *DCW1*-FLAG alleles, of which we chose 10 for further analysis. All 10 *DCW1*-FLAG alleles that we created proved to have wild-type levels of function when assayed for growth at elevated temperatures.

Monitoring the distribution of Dcw1 in fractionated cell lysates by Western blot illustrated that the available FLAG antibodies were more sensitive than antibodies directly raised against a peptide in Dcw1. The majority of alleles had a subcellular distribution similar to the wild-type protein—primarily in the membrane fraction, consistent with its proposed function (Kitagaki *et al.* 2002). Two alleles, Dcw1-$FLAG_5$ and Dcw1-$FLAG_6$, were not detectable in any cell fraction in the α-FLAG Western blots. We suspect that the FLAG tag, located in the signal sequence in these two alleles, may have been cleaved off

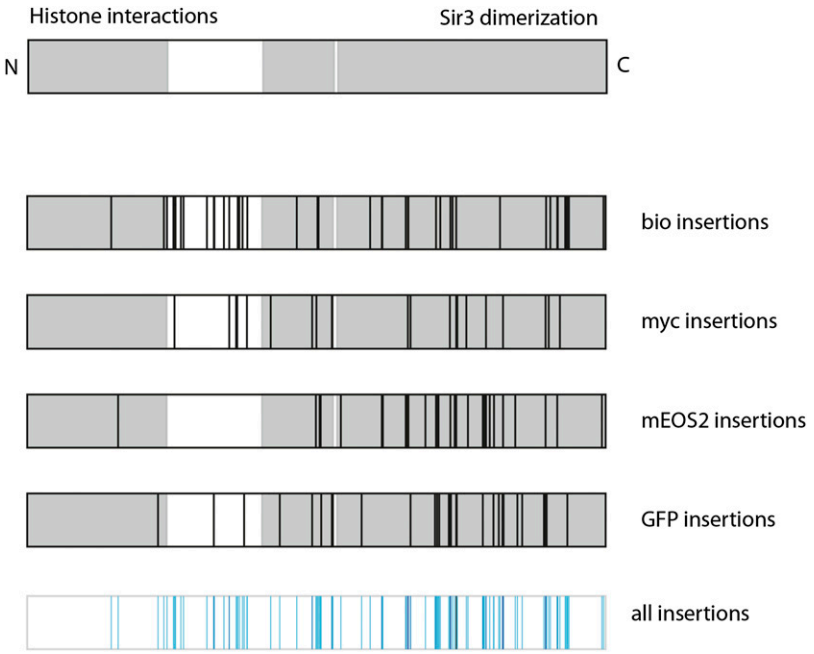

Figure 6 Distribution of functional tagged alleles of *C. glabrata SIR3*. (Top) Linear representation of the *C. glabrata* Sir3 protein (1088 residues). The gray areas are predicted globular domains (by ELM), which have been found to have roles in associating with histones (N-terminal domain) and Sir3 dimerization (C-terminal domain). In the schematics below, a black line is located at the position for each unique insertion site identified from among the functional Sir3-bio, Sir3-myc, Sir3-mEOS2, and Sir3-GFP pools. (Bottom) Schematic overlays all sites at which a functional tagged Sir3 was identified. The color of the line indicates the number of mutagenesis pools in which the particular insertion site was identified; light blue indicates that the site was identified in only one pool; black indicates that it was found in all four mutagenesis pools.

during normal protein trafficking. In the case of Dcw1-$FLAG_{11}$, the protein was detected in Western blots, but appeared smaller than the expected size (Figure 5). This apparent size difference could be due to truncation or incomplete post-translational glycosylation. We note that introduction of internal epitopes may still alter protein function. However, if multiple functional tagged alleles have the same phenotype and show the same subcellular distribution —as we observed for 7 of 10 alleles—this increases confidence that they reflect wild-type function.

We also mutagenized the *C. glabrata SIR3* gene and isolated 126 functional, unique internally tagged alleles (four different tags). By Western analysis, GFP internally tagged proteins were expressed at similar levels, and all showed a punctate distribution similar to that documented for Sir3 in *S. cerevisiae*. Since the tranposon insertion is random and generates many nonfunctional insertions, the utility of the method depends on a robust screen for functional alleles. This can be an assay for viability (as for *DCW1*). For Sir3, the assay for function was not viability but growth in the presence of 5-FOA as a measure of transcriptional silencing of a subtelomeric *URA3* gene. For *DCW1* and *SIR3*, ~5% of insertions were functional, suggesting that, for some proteins at least, functional alleles could also be identified by manual screening of a relatively small number of insertions.

For *DCW1*, the functional alleles had FLAG insertions clustered in a few regions of the gene, including in the central globular glycosyl hydrolase domain (identified by protein BLAST and ELM motif finding) (Kitagaki *et al.* 2002). For *SIR3* as well, functional insertions for all four tags were found across the ORF, including insertions in N- and C-terminal predicted globular domains. It would have been difficult to predict that these regions would tolerate insertions solely based on bioinformatic modeling of the protein structure. Unlike other epitope-tagging systems (Khmelinskii *et al.* 2011; Ramsden *et al.* 2011), this Tn7-based system requires no prior knowledge of permissive sites to introduce tags in the POI. The efficient *in vitro* mutagenesis and cloning steps allow researchers to create and recover an unbiased set of tagged versions of a target gene, which can then be screened for function.

The system of Tn7-based epitope tagging that we describe should be broadly useful. We have described the application to epitope tagging of an essential yeast gene. However, several aspects of the methodology make it suitable to tagging of proteins in virtually any system. First, once Gateway vectors are constructed, mutagenesis, cloning, and screening (in yeast) takes 2–3 weeks total. Second, the mutagenesis is carried out *in vitro*, and Tn7 has minimal insertion sequence bias. Third, the insertions in the ORF are isolated without regard to function of those insertions. Fourth, the library of tagged proteins is generated in final form in *E. coli*, obviating the need to remove transposon sequences in the cell type where functional screening is done. Rather, the final library can simply be introduced into the appropriate cell type and screened directly.

Since a major use of transposon epitope-tagging methods is to generate fully functional tagged versions of the POI, it is imperative that the screening be done only on mutagenized ORFs since any untagged wild-type alleles in the pool would pass the functional screen. In the highly useful Tn5-based system, TAGIT, the selection for tagged alleles is by presence of the marker gene carried in the transposon. This allows introduction of the mutagenized library into the cell where screening is carried out; at this stage the transposon sequences are removed by expression in the cell of Cre recombinase. Our

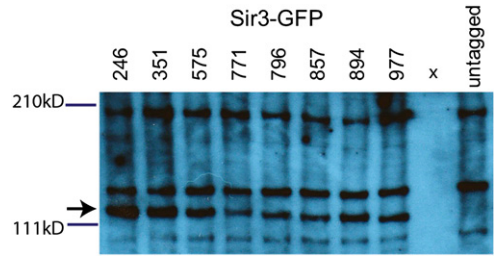

Figure 7 Western blot analysis of Sir3-GFP in *C. glabrata*. An anti-GFP immunoblot of *C. glabrata* strains carrying different Sir3-GFP alleles. Wells are labeled with the amino acid site of the GFP insertion within Sir3; "x" indicates an empty lane. The antibody cross-reacted with *C. glabrata* whole-cell lysates, as shown by the banding see in the untagged strain. The band representing Sir3-GFP is marked with an arrow.

method provides an alternative to this approach. Insertions in the ORF are efficiently isolated from nonmutagenized ORFs by Gateway-mediated recombination. This step is highly efficient and can be carried out in batches, permitting construction of a library of tagged ORFs comprising tens or hundreds of thousands of independent insertion events. Subsequent efficient restriction digestions of the pool to remove the left and right ends of the transposon result in a complex pool of tagged ORFs in an appropriate expression vector. This library can then be directly screened for function as dictated by function of the POI, without the need to express recombinase enzymes in the target cell. As a caveat, we note that we recovered a small number of functional expression plasmids where the epitope had inserted between the Gateway recombination sites and the ORF (data not shown); it appears that the attL1 and attL2 sites are somewhat permissive for insertions. This leaves the target ORF unmutagenized, and our current plasmid architecture makes it impossible to eliminate these extraneous plasmids from the pool.

An advantage of this Tn7-based method over other transposon systems is the very limited amount of transposon-derived sequence left in the insertion site. In our method, the *Fse*I and *Pme*I sites are engineered to be present within three bases of the transposon end, and following excision, there are only 22 nucleotides (essentially the *Fse*I and *Pme*I sites) in addition to the epitope tag present at the insertion site. This increases, we would argue, the likelihood of identifying functional epitope insertions since it limits the amount of extraneous sequence shared by all insertions to a few amino acids.

Finally, the method in principle could be used for analysis of proteins in many different organisms. The mutagenesis of the cloned gene is done *in vitro*, and mobilization into the expression vector exploits Gateway recombination. This can be adapted most easily to any system for which Gateway-modified expression vectors exist. We suggest that the method will be useful in a range of systems, including model eukaryotic systems like *S. cerevisiae* as well as mammalian systems.

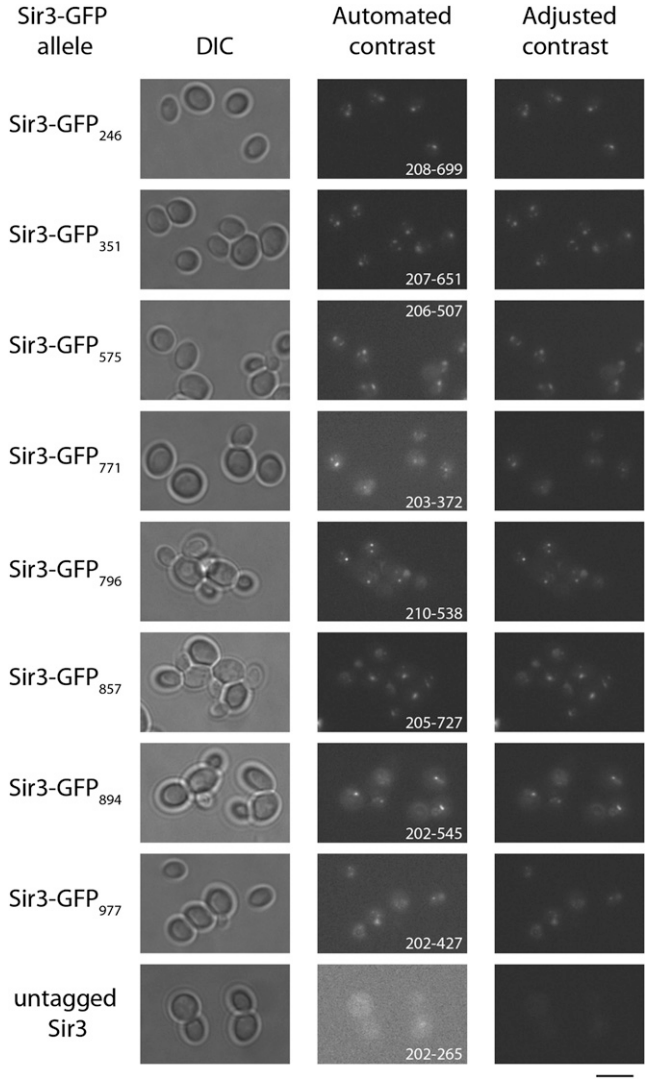

Figure 8 Imaging of live *C. glabrata* strains carrying various Sir3-GFP alleles. Eight strains carrying different versions of Sir3-GFP, as well as an untagged Sir3 control strain, were grown to stationary phase in liquid culture. Live cells were imaged using differential interference contrast bright-field and fluorescence microscopy. GFP fluorescence is shown with both the contrast settings as captured (automated contrast) and adjusted so each strain has the same contrast levels (adjusted contrast). The maximum and minimum contrast settings for each strain are listed on the automated contrast panel. The adjusted contrast panels all have a minimum of 202 and a maximum of 727. Bar, 5 μm.

Acknowledgments

We thank Matthew Frieman for creating the *dcw1*Δ*dfg5*Δ pCU-*DCW1* strain BY240. We thank Irene Castaño for strain CGM293, Yuxia Ren for her assistance cloning the Tn7 donor vectors, and Jie Xiao for the gift of pET28-ftsZ-mEOS. R.E.Z. is supported by an American Cancer Society Postdoctoral Fellowship (PF-09-194-01-MPC). This work was supported by National Institutes of Health grants R01A1046223 (to B.P.C.) and RO1GM076425 (to N.L.C.). N.L.C. is an Investigator of the Howard Hughes Medical Institute.

Literature Cited

Arnau, J., C. Lauritzen, G. E. Petersen, and J. Pedersen, 2006 Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr. Purif. 48: 1–13.

Bell, M. R., M. J. Engleka, A. Malik, and J. E. Strickler, 2013 To fuse or not to fuse: What is your purpose? Protein Sci. 22: 1466–1477.

Biery, M. C., F. J. Stewart, A. E. Stellwagen, E. A. Raleigh, and N. L. Craig, 2000 A simple in vitro Tn7-based transposition system with low target site selectivity for genome and gene analysis. Nucleic Acids Res. 28: 1067–1077.

Castano, I., R. Kaur, S. Pan, R. Cregg, L. Penas Ade *et al.*, 2003 Tn7-based genome-wide random insertional mutagenesis of Candida glabrata. Genome Res. 13: 905–915.

Choi, K. Y., Y. Li, R. Sarnovsky, and N. L. Craig, 2013 Direct interaction between the TnsA and TnsB subunits controls the heteromeric Tn7 transposase. Proc. Natl. Acad. Sci. USA 110: E2038–E2045.

Cockell, M., F. Palladino, T. Laroche, G. Kyrion, C. Liu *et al.*, 1995 The carboxy termini of Sir4 and Rap1 affect Sir3 localization: evidence for a multicomponent complex required for yeast telomeric silencing. J. Cell Biol. 129: 909–924.

Cormack, B. P., and S. Falkow, 1999 Efficient homologous and illegitimate recombination in the opportunistic yeast pathogen *Candida glabrata*. Genetics 151: 979–987.

De Las Penas, A., S. J. Pan, I. Castano, J. Alder, R. Cregg *et al.*, 2003 Virulence-related surface glycoproteins in the yeast pathogen Candida glabrata are encoded in subtelomeric clusters and subject to RAP1- and SIR-dependent transcriptional silencing. Genes Dev. 17: 2245–2258.

Devine, S. E., and J. D. Boeke, 1994 Efficient integration of artificial transposons into plasmid targets in vitro: a useful tool for DNA mapping, sequencing and genetic analysis. Nucleic Acids Res. 22: 3765–3772.

Dujon, B., D. Sherman, G. Fischer, P. Durrens, S. Casaregola *et al.*, 2004 Genome evolution in yeasts. Nature 430: 35–44.

Frieman, M. B., and B. P. Cormack, 2004 Multiple sequence signals determine the distribution of glycosylphosphatidylinositol proteins between the plasma membrane and cell wall in Saccharomyces cerevisiae. Microbiology 150: 3105–3114.

Gamas, P., and N. L. Craig, 1992 Purification and characterization of TnsC, a Tn7 transposition protein that binds ATP and DNA. Nucleic Acids Res. 20: 2525–2532.

Gonzalez, M., N. Goddard, C. Hicks, R. Ovalle, J. M. Rauceo *et al.*, 2010 A screen for deficiencies in GPI-anchorage of wall glycoproteins in yeast. Yeast 27: 583–596.

Green, B., C. Bouchier, C. Fairhead, N. L. Craig, and B. P. Cormack, 2012 Insertion site preference of Mu, Tn5, and Tn7 transposons. Mob. DNA 3: 3.

Gregory, J. A., E. C. Becker, J. Jung, I. Tuwatananurak, and K. Pogliano, 2010 Transposon assisted gene insertion technology (TAGIT): a tool for generating fluorescent fusion proteins. PLoS ONE 5: e8731.

Haldimann, A., M. K. Prahalad, S. L. Fisher, S. K. Kim, C. T. Walsh *et al.*, 1996 Altered recognition mutants of the response regulator PhoB: a new genetic strategy for studying protein-protein interactions. Proc. Natl. Acad. Sci. USA 93: 14361–14366.

Hill, J., K. A. Donald, and D. E. Griffiths, 1991 DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res. 19: 5791.

Hoffman, C. S., 2001 Preparation of yeast DNA. Curr. Protoc. Mol. Biol. Chapter 13: Unit13 (13.11-1-13.11.4).

Khmelinskii, A., M. Meurer, N. Duishoev, N. Delhomme, and M. Knop, 2011 Seamless gene tagging by endonuclease-driven homologous recombination. PLoS ONE 6: e23794.

Kitagaki, H., H. Wu, H. Shimoi, and K. Ito, 2002 Two homologous genes, DCW1 (YKL046c) and DFG5, are essential for cell growth and encode glycosylphosphatidylinositol (GPI)-anchored membrane proteins required for cell wall biogenesis in Saccharomyces cerevisiae. Mol. Microbiol. 46: 1011–1022.

Kitagaki, H., K. Ito, and H. Shimoi, 2004 A temperature-sensitive dcw1 mutant of Saccharomyces cerevisiae is cell cycle arrested with small buds which have aberrant cell walls. Eukaryot. Cell 3: 1297–1306.

Kolter, R., M. Inuzuka, and D. R. Helinski, 1978 Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. Cell 15: 1199–1208.

Kumar, A., S. Agarwal, J. A. Heyman, S. Matson, M. Heidtman *et al.*, 2002 Subcellular localization of the yeast proteome. Genes Dev. 16: 707–719.

Kumar, A., M. Seringhaus, M. C. Biery, R. J. Sarnovsky, L. Umansky *et al.*, 2004 Large-scale mutagenesis of the yeast genome using a Tn7-derived multipurpose transposon. Genome Res. 14: 1975–1986.

Maddi, A., C. Fu, and S. J. Free, 2012 The Neurospora crassa dfg5 and dcw1 genes encode alpha-1,6-mannanases that function in the incorporation of glycoproteins into the cell wall. PLoS ONE 7: e38872.

Manoil, C., and B. Traxler, 2000 Insertion of in-frame sequence tags into proteins using transposons. Methods 20: 55–61.

Merkulov, G. V., and J. D. Boeke, 1998 Libraries of green fluorescent protein fusions generated by transposition in vitro. Gene 222: 213–222.

Metcalf, W. W., W. Jiang, L. L. Daniels, S. K. Kim, A. Haldimann *et al.*, 1996 Conditionally replicative and conjugative plasmids carrying lacZ alpha for cloning, mutagenesis, and allele replacement in bacteria. Plasmid 35: 1–13.

Mumberg, D., R. Muller, and M. Funk, 1995 Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene 156: 119–122.

Osawa, M., and H. P. Erickson, 2005 Probing the domain structure of FtsZ by random truncation and insertion of GFP. Microbiology 151: 4033–4043.

Penfold, R. J., and J. M. Pemberton, 1992 An improved suicide vector for construction of chromosomal insertion mutations in bacteria. Gene 118: 145–146.

Ramsden, R., L. Arms, T. N. Davis, and E. G. Muller, 2011 An intein with genetically selectable markers provides a new approach to internally label proteins with GFP. BMC Biotechnol. 11: 71.

Ross-Macdonald, P., P. S. Coelho, T. Roemer, S. Agarwal, A. Kumar *et al.*, 1999 Large-scale analysis of the yeast genome by transposon tagging and gene disruption. Nature 402: 413–418.

Seringhaus, M., A. Kumar, J. Hartigan, M. Snyder, and M. Gerstein, 2006 Genomic analysis of insertion behavior and target specificity of mini-Tn7 and Tn3 transposons in Saccharomyces cerevisiae. Nucleic Acids Res. 34: e57.

Sheridan, D. L., C. H. Berlot, A. Robert, F. M. Inglis, K. B. Jakobsdottir *et al.*, 2002 A new way to rapidly create functional, fluorescent fusion proteins: random insertion of GFP with an in vitro transposition reaction. BMC Neurosci. 3: 7.

Spreghini, E., D. A. Davis, R. Subaran, M. Kim, and A. P. Mitchell, 2003 Roles of Candida albicans Dfg5p and Dcw1p cell surface proteins in growth and hypha formation. Eukaryot. Cell 2: 746–755.

Ubersax, J. A., E. L. Woodbury, P. N. Quang, M. Paraz, J. D. Blethrow *et al.*, 2003 Targets of the cyclin-dependent kinase Cdk1. Nature 425: 859–864.

Winzeler, E. A., D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson *et al.*, 1999 Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science 285: 901–906.

Young, C. L., Z. T. Britton, and A. S. Robinson, 2012 Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol. J. 7: 620–634.

Zordan, R. E., Y. Ren, S. J. Pan, G. Rotondo, A. De Las Penas *et al.*, 2013 Expression plasmids for use in Candida glabrata. G3 (Bethesda) 3: 1675–1686.

Communicating editor: O. Cohen-Fix

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.169482/-/DC1

Avoiding the Ends: Internal Epitope Tagging of Proteins Using Transposon Tn7

Rebecca E. Zordan, Brian J. Beliveau, Jonathan A. Trow, Nancy L. Craig, and Brendan P. Cormack

DOI: 10.1534/genetics.114.169482

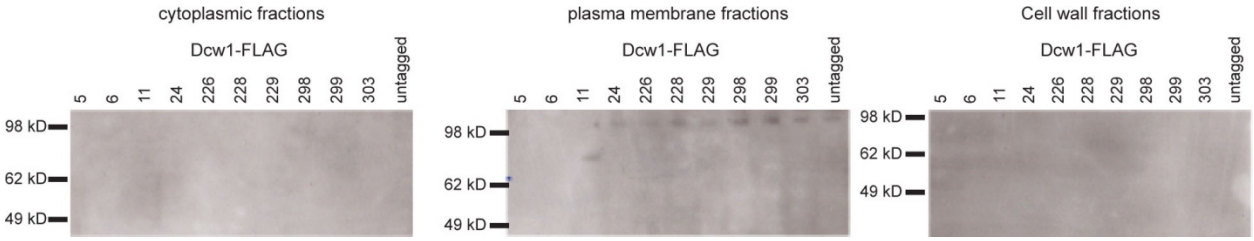

Figure S1. Western blot analysis of Dcw1-FLAG strains with α-DCW1 antibody

The same cell fractions used in Figure 5 were run on new SDS-PAGE gels and probed with an α-Dcw1 antibody. Signal was detected using chemiluminescence, and film was exposed for 45 minutes. Material from $7.5x10^5$ cells is loaded in each strain in each fraction. Wells are labeled with the amino acid site of the FLAG tag within Dcw1. We note the apparent size of Dcw1 in the plasma membrane is larger than what we observed with α-FLAG antibodies, but the overall pattern is consistent with the α-FLAG Westerns shown in Figure 5.

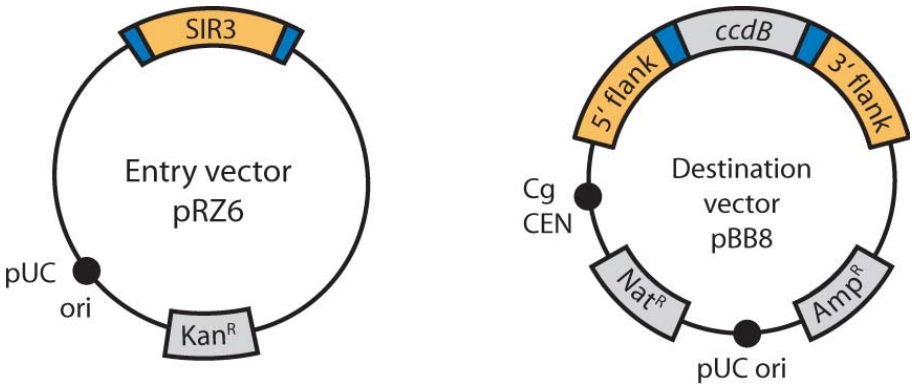

Figure S2. Gateway constructs used for mutagenesis of *C. glabrata SIR3*
These schematic drawings represent the Gateway entry vector and destination vector used during the mutagenesis of *C. glabrata SIR3*, as represented in Figure 2 in the manuscript.

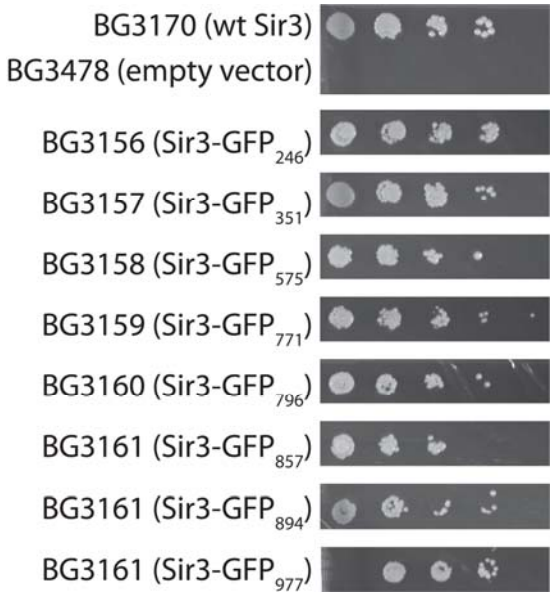

Figure S3. Sir3-GFP alleles complement for growth on 5-FOA

C. glabrata strains carrying various Sir3-GFP alleles were grown overnight in YPD+Nat media in 96-well plates. Yeast were diluted to OD_{600}=2 and placed in the left-most well of a new 96-well plate. Ten-fold serial dilutions were made in sterile PBS, and the cells were transferred to a 5-FOA plate using a multi-pin tool. The plate was grown at 30°C. We note that the lack of growth for BG3161 in the left-most (densest) dilution is likely due to uneven placement of the pin tool onto the 5-FOA plate.

Table S1. Plasmids used in this study

Plasmid	Description	*E. coli* marker	Yeast marker
Tn7 donor plasmids			
pRZ49	Tn7-mEOS2 donor vector	Kan	n/a
pRZ98	Tn7-biotin-donor vector	Kan	n/a
pRZ99	Tn7-6His donor vector	Kan	n/a
pRZ101	Tn7-FLAG donor vector	Kan	n/a
pRZ102	Tn7-HA donor vector	Kan	n/a
pRZ103	Tn7-myc donor vector	Kan	n/a
pRZ106	Tn7-GFP donor vector	Kan	n/a
DCW1 plasmids			
-	pCU-*DCW1*	Amp	*URA3*
-	*DCW1* entry vector	Kan	n/a
pRZ159	empty *DCW1* destination vector	Amp	*HIS1*
pRZ160	wild type *DCW1* expression vector	Amp	*HIS1*
pRZ165	Dcw1-$FLAG_5$	Amp	*HIS1*
pRZ166	Dcw1-$FLAG_6$	Amp	*HIS1*
pRZ172	Dcw1-$FLAG_{11}$	Amp	*HIS1*
pRZ173	Dcw1-$FLAG_{24}$	Amp	*HIS1*
pRZ174	Dcw1-$FLAG_{226}$	Amp	*HIS1*
pRZ175	Dcw1-$FLAG_{228}$	Amp	*HIS1*
pRZ167	Dcw1-$FLAG_{229}$	Amp	*HIS1*
pRZ168	Dcw1-$FLAG_{298}$	Amp	*HIS1*
pRZ176	Dcw1-$FLAG_{299}$	Amp	*HIS1*
pRZ177	Dcw1-$FLAG_{303}$	Amp	*HIS1*
pBC715	Dcw1-HA_{26}	Amp	*HIS1*
SIR3 plasmids			
pRZ6	*SIR3* entry vector	Kan	n/a
pBB8	empty *SIR3* destination vector	Amp	Nat^R
pRZ47	wild type *SIR3* expression vector	Amp	Nat^R
pRZ120	Sir3-GFP_{246}	Amp	Nat^R
pRZ121	Sir3-GFP_{351}	Amp	Nat^R
pRZ122	Sir3-GFP_{575}	Amp	Nat^R
pRZ124	Sir3-GFP_{771}	Amp	Nat^R
pRZ127	Sir3-GFP_{796}	Amp	Nat^R
pRZ130	Sir3-GFP_{857}	Amp	Nat^R
pRZ131	Sir3-GFP_{894}	Amp	Nat^R
pRZ133	Sir3-GFP_{977}	Amp	Nat^R

Table S2. Primers used in this study

Oligo #	Name	Sequence (5'-3')
Tn7-tag donor vector construction		
3201	Tn7L-FseI-Bam-for	aggactacggatcctgtggccggccAATAAAGTCTTAAACTGAACAAA
3202	Tn7L-AscI-rev	gacctgacggcgcgccGTCGACCCCACGCCCCTCTTTAAT
4724	mEOS_for	aggcgcgccggccggccTGGATCCGCTGGCTCCGCTGCTGGTTCTGGC GAATTCATGAGT
4725	mEOS_rev	gctctagagtttaaacTAAATTCTCCAGATCCTGCAGCAGATCCTGCAGAGCCTCGTCT GGCATTGTCAGGCAATCCAGAATGAG
5300	L1-GFP-f	atttaggatccgctggctccgctgctggttctggcATGTCTAAAGGTGAAGAATTATTCACTGG TG
5301	L2-GFP-r	ataagtttaaactagctcctcctgcagcagatcctgcagagccTTTGTACAATTCATCCATACCAT GGGTAATAC
DCW1-Yiplac211 knockout construct		
1975	*DCW1* 3'flank rev SphI	acatgcatgcAGGAAACCATGTAAGCGATGAATAT
1976	*DCW1* 3'flank for KpnI	ggggtaccTGCAGAACTTATGAAAGCTTAACATTT
1977	*DCW1* 5'flank rev KpnI	ggggtaccTTTTATGTGTTCGTTTTTAAAACAGAC
1978	*DCW1* 5'flank for HindIII	ccccaagcttAGATGAACTTGAACTTAAGATGATC
Amplify across *DCW1* knockout region		
2306	*DCW1* 5' check	TCGTTTAAATTCAATTGGAACTGTA
2307	*DCW1* 3' check	TTCAAACAAAATTCGTTCGATATTA
Verify *DCW1*-Yiplac211 integrants		
1504	Yiplac backbone	TATGTTGTGTGGAATTGTGAGCGG
1505	URA check	GCGATTAAGTTGGGTAACGCCAGG
Verify loopout of *DCW1*-Yiplac211 construct		
1778	*DCW1* 5' check	ACCTTTCCAGGACATATAAT
1779	*DCW1* 3' check	ACACATATGAACAAAGGTCT
pCU-*DCW1* construction		
1629	3ecoYKL046	ccggaattcTCAAAAGACTAACCACAGAGCACATG
1630	5bamYKL046	cgcgggatccATGCTAGTAAATAAAGTGATAGGGT
DCW1 destination vector		
6333	*DCW1* promoter - for	atagagctcTTCTTCTCCTTATTGTGCTTTACC
6334	*DCW1* promoter - rev	attctagaTTTTATGTGTTCGTTTTTAAAACAGACTG
Determine position of FLAG insertion.		
2766	*DCW1* promoter - for	GATGATCATAGGTACTCTTTGTATAATGGGC
6244	linker 2 - rev	ATTAGTTTAAACTAGCTCCTCCTGCA
5160	linker 1 - for	CTCCGCTGCTGGTTCTGG
4032	M13F (-21)	GTAAAACGACGGCCAGT
5626	Tn7L - rev	GATCTATTTTGTTCAGTTTAAGACTTTATTG
Sir3 entry vector construction		
4470	SIR3b1	ggggacaagtttgtacaaaaaagcaggctaaaaccATGGCTGAGCTTATAAAAGACCTG
4484	SIR3b2	ggggaccactttgtacaagaaagctgggtCTATTCGGTGAGACACGATTGGAT

Sir3 destination vector construction			
	3220	Sir3_IP_SacI_F	gtacctatgagctcGAACGGTGCCAGACACACCAGCCC
	3221	Sir3_IP_XbaI_R	tgaccatatctagaCCTCTTACTTAATCCGAAACCTTC
	3222	Sir3_UTR_XhoI_F	caatgcacactcgagAAAAGCTTTCATCTTCTTTTCTTGATTCTCCTC
	3223	Sir3_UTR_KpnI_R	catgaccatggtaccAAGACGGCTCCATCACTAAAGTGC
Determine position of epitope insertions within *SIR3* using colony PCR and sequencing			
	5160	L1 for	CTCCGCTGCTGGTTCTGG
	5161	L2 rev	CTCCTCCTGCAGCAGATCCT
	5602	SIR3IP_3'for	CTGGGAAGGTTTCGGATTAAGTAAGAGG
	5605	SIR3utr_5'rev	GTATTAGTAGAGGAGAATCAAGAAAAGAAGATGAAAG
	5627	Linker2+myc	GATCCTGCAGAGCCTTCATTGAG
	5738	Bio_L2-rev	GATCCTGCAGAGCCTTCATGCC
	5739	GFP 5' rev	AGGTCAATTTACCGTAAGTAGCATCAC
	5740	GFP 3' for	TTATCCACTCAATCTGCCTTATCCA
	5741	mEOS 5' rev	CGAATACCCTGTTGCCGTAATGGA
	5742	mEOS 3' for	ACCGATGTGACTTCAGAACTACTTACAAAG

Bases that anneal to the template are shown in capital letters; restriction sites, linkers, and Gateway recombination sites added to the primers are shown in lowercase letters.

Table S3. *S. cerevisiae* strains used in this study

Strain	Genotype	Parent	Source
BY4741	MAT**a** *his3*Δ1 *leu2*Δ0 *met15*Δ0 *ura3*Δ0		[Brachmann, 1998 #76]
BY4742	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0		[Brachmann, 1998 #76]
*dfg5*Δ	MAT**a** *his3*Δ1 *leu2*Δ0 *met15*Δ0 *ura3*Δ0 *dfg5*::Kan^R	BY4741	[Winzeler, 1999 #79]
*dcw1*Δ	MAT**a** *his3*Δ1 *leu2*Δ0 *met15*Δ0 *ura3*Δ0 *dcw1*(YKL046C)::Kan^R	BY4741	[Winzeler, 1999 #79]
BY240	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pCU-*DCW1*	BY4742	this work
BY965	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pCU-*DCW1* pRZ159	BY240	this work
BY966	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ160	BY240	this work
BY893	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ165	BY240	this work
BY895	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ166	BY240	this work
BY974	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ172	BY240	this work
BY975	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ173	BY240	this work
BY976	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ174	BY240	this work
BY977	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ175	BY240	this work
BY897	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ167	BY240	this work
BY899	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ168	BY240	this work
BY978	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ176	BY240	this work
BY979	MATα *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *dfg5*::g418 *dcw1*Δ pRZ177	BY240	this work

Table S4. *C. glabrata* strains used in this study

Strain	Genotype	Parent	Source
CGM293	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4*	CGM693	[Rosas-Hernandez, 2008 #105]
BG3156	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ120	CGM293	this work
BG3157	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ121	CGM293	this work
BG3158	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ122	CGM293	this work
BG3159	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ124	CGM293	this work
BG3160	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ127	CGM293	this work
BG3161	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ130	CGM293	this work
BG3162	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ131	CGM293	this work
BG3163	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ133	CGM293	this work
BG3170	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pRZ47	CGM293	this work
BG3478	*ura3Δ*::Tn*903* G418^R, *sir3Δ::hph* (pIC120/*Kpn* I-*Hpa* I+T4 DNA polymerase). *URA3* at unique region between *EPA5* and *EPA4* pBB8	CGM293	this work

Table S5. **Drug resistance of isolates from xpC culture, as assessed from replica plating.**

Source plate	$Tmp^RKan^RCar^R$ colonies	Tmp^RCar^R colonies	Kan^RCar^R colonies	Only Car^R colonies	Total Car^R colonies	Tmp^RCar^R/ (total Car^R) correction
1	2	14	18	150	184	7.6%
2	5	9	24	119	157	5.7%
3	2	5	24	166	197	2.5%
4	1	8	23	158	190	4.2%

"Total Car^R colonies" is the sum of all colonies on the previous 4 columns (and matches the total number of colonies on the original LB+Car source plate. Tmp^RCar^R /(total Car^R) correction indicates the percentage of all Car^R colonies that are in fact Tmp^R Car^R. This represents the % of the xpC pool that bears a mutagenized *DCW1*-FLAG expression vector, and was used to correct the xpC pool size shown in Table 1 of the main paper.

Table S6. Size of *SIR3* mutagenesis pools during selection and sequencing

Colony counts, as plated after recovery during transformation, prior to drug selection in pool. Cells were diluted appropriately to prevent a lawn of growth and plated onto media as indicated. Numbers in the table are calculated to represent the number of colonies in the full transformation resistant to the given drug. nd = not determined. Fold coverage was calculated by comparing the number of colonies from the indicated drug selection to the size of the DNA available for mutagenesis (Supplemental Note 1). * The Sir3-myc pool was treated differently than other pools. After transformation of the Sir3-myc pool into *C. glabrata*, the pooled culture was outgrown in YPD+NAT for ~24 hours prior to plating on YPD+NAT. All other tagged Sir3 pools were plated directly onto YPD+NAT plates following transformation and recovery.

		Tn used to mutagenize *C. glabrata SIR3*							
Step of processing	**Fold coverage based on which drugs?**	**myc**		**bio**		**mEOS2**		**GFP**	
		# colonies	**fold coverage**	**# colonies**	**fold coverage**	**# colonies**	**fold coverage**	**# colonies**	**fold coverage**
Mutagenesis of entry vector	Tmp Kan	5.4E+05	67x	2.9E+05	36x	3.2E+05	39x	6.46E+05	**79.99**
xpC = Gateway LR to create expression pool	Tmp Car	6.0E+06	920x	6.0E+06	920x	6.0E+06	920x	6.00E+06	920x
xpT	Tmp Car	5.7E+07	8700x	8.2E+07	13000x	1.8E+07	2800x	2.70E+07	4100x
xpT-Tn7L	Tmp Car	1.2E+07	1800x	3.6E+07	5500x	4.1E+07	6200x	2.34E+07	3600x
xpT-Tn7L-Tn7R	Car	1.6E+07	2400x	5.8E+07	8900x	6.2E+07	9500x	3.50E+07	5400x
Transformation into *C. glabrata*	Nat	7992*	1.2x	2172	0.33x	2432	0.37x	1895	0.29x
Screening for functional isolates	5-FOA	906	n/a	190	n/a	162	n/a	105	n/a
% functional clones		11%	n/a	8.7%	n/a	6.7%	n/a	5.5%	n/a
Sent for sequencing		95		63		52		48	
# mappable sequences		90		53		47		48	
# isolates with inframe insertions in SIR3 orf		88		52		46		47	
# in-frame unique insertions sites		21		42		33		30	
Max # insertions at a single site		27		4		6		6	

Insertion site (amino acid)	Epitope tag			
	myc	bio	mEOS2	GFP
158		1		
171			1	
246				5
257		1		
263		1		
275		1		
277	1			
278		1		
279		1		
289		1		
294		1		
338		1		
351		1		2
352		1		
370		4		
380	1	1		
393	3			
395	15			
396		1		
399		1		
405		1		
408				1
413	3			
414		2		
458	5			
475				1
507		2		
536	1			2
543			1	
544	1			
546		1		
548		1		
550			1	
551			1	
552			2	
553				1
573	2			1
574	1			
575				1
590			2	
629				1
645		1		
667		1	1	
668		1		
669			1	
712		1	1	
713			1	
716	2	1	1	
717		1	1	
721	1			1
749			1	
767				1

Table S7. Insertion sites of select in-frame complementing tagged Sir3-alleles

Complementing clones from the mutagenized Sir3 pools were analyzed with colony PCR to qualitatively determine the location of the epitope tag within the *SIR3* ORF. PCR products from a subset of these complementing clones were sequenced and the precise location of the tag insertion was identified. Each of the 107 unique locations within Sir3 where a tag was found in a functional allele is listed in the "insertion site" column, based on the amino acid of insertion. The final two insertion sites listed are downstream of the Sir3 STOP codon. The remaining columns list the number of complementing clones from each mutagenized pool found at a given insertion location. In total, the epitope tag was located in frame in 65 Sir3-myc, 52 Sir3-bio, 46 Sir3-mEOS2, and 47 Sir3-GFP complementing clones.

Insertion Site (Amino acid)	Epitope tag			
	myc	bio	mEOS2	GFP
769		1	4	
771				1
772			1	
773			1	
775				1
777		2		
793				1
795	2			
796		3	2	4
797				1
800		1		
804			1	
807	7	3	6	6
808				1
809	2			
826	27			
829			1	
857			1	1
858			2	
859			1	
860			1	
863	1		1	
870			1	
877				1
878			1	
887				1
889		1		
894			2	1
895	2		1	1
896				1
918			1	
922				1
931				1
972				1
975	7		1	1
976		1		
977				4
981	3			
984		1		
997		1	1	
998		1		
1002	1			
1012		1		
1013		1		
1015		1		
1016				1
1017		1		
1019		1		
1081			1	
1084		1		
_+6 bp	1			
+17 bp	1			

Table S7 (continued). Insertion sites of select in-frame complementing tagged Sir3-alleles

Complementing clones from the mutagenized Sir3 pools were analyzed with colony PCR to qualitatively determine the location of the epitope tag within the *SIR3* ORF. PCR products from a subset of these complementing clones were sequenced and the precise location of the tag insertion was identified. Each of the 107 unique locations within Sir3 where a tag was found in a functional allele is listed in the "insertion site" column, based on the amino acid of insertion. The final two insertion sites listed are downstream of the Sir3 STOP codon. The remaining columns list the number of complementing clones from each mutagenized pool found at a given insertion location. In total, the epitope tag was located in frame in 95 Sir3-myc, 52 Sir3-bio, 46 Sir3-mEOS2, and 47 Sir3-GFP complementing clones.

File S1

Protocol

In vitro Transposon Mutagenesis for Introduction of Internal Epitope Tags

REQUIRED REAGENTS

Purified TnsA, TnsB, and TnsC^{A255V} enzymes [3, 4]

TnsA stock = 150ng/μl in Storage Buffer A

TnsB stock = 200ng/μl in Storage Buffer B

TnsC ^{A255V} stock = 500ng/μl in Storage buffer C

Tn7 donor vector: pRZ101 (Tn7-FLAG donor vector). Dilute to 25ng/μl in 10mM Tris pH8.0

Entry vector of target gene: *DCW1* entry vector. Dilute to 50ng/μl in 10mM Tris pH8.0

Invitrogen MegaX DH10B T1R electrocompetent *E. coli* cells

300mM Magnesium Acetate (MgOAc)

phenol:chloroform:IAA (25:24:1) (Amresco 0883-100ml)

Chloroform

3M Sodium Acetate (NaOAc)

100% ethanol (EtOH)

70% EtOH

FseI (New England Biolabs R0588S)

PmeI (New England Biolabs R0560S)

ApaLI (New England Biolabs R0507S)

T4 DNA ligase and buffer (New England Biolabs M0202S)

Gateway LR clonase II (Life Techonologies 11791-100)

Qiagen Hi-Speed MidiPrep kit (Qiagen 12643)

MegaX DH10B T1R Electrocomptent Cells (Life Technologies C6400-03)

LB media

Oxoid Isosensitest media (Iso) (agar: OXCM0471B, liquid broth: OXCM0473B)

Carbenicillin (100mg/ml = 1000x stock)

Kanamycin (30mg/ml = 1000x stock)

Trimethoprim (5mg/ml in DMSO = 500x stock)

ddH_2O = doubly distilled (MilliQ) H_2O

BUFFERS

Use sterile-filtered MilliQ water to make all buffers. If possible, make all buffers in plastic containers; residual detergent on glassware may inhibit the transposition reaction.

Standard buffers

20mM ATP in 125mM Tris pH 7.5 (store at -20°C for at most 1 month)
20mM DTT in 125mM Tris pH 7.5 (store at -20°C for at most 1 month)
100mM ATP in 250mM HEPES (pH8.0) – for making TnsC storage buffer
1M DTT in 150mM HEPES (pH8.0) – for making TnsA and TnsC storage buffers
1M DTT in 1M Tris pH 7.5 – for making TnsB storage buffer
10mM Tris - for elution of mutagenized plasmid pools from midiprep kit
50% glycerol

Tns storage buffers

Store at -20°C for at most 6 months. We do not recommend refreezing and rethawing of the storage buffers, so store them in small aliquots and discard after use.

Storage buffer A
25mM HEPES pH8.0
150mM NaCl
1mM EDTA
1mM DTT (in HEPES)
10% glycerol

Storage buffer B
25mM TrisHCl pH8.0
500mM NaCl
1mM EDTA
1mM DTT (in Tris)
25% glycerol

Storage buffer C
25mM HEPES pH8.0
1M NaCl
0.1mM EDTA
2.5mM DTT (in HEPES)
1mM ATP
10mM $MgCl_2$
10% glycerol

***In vitro* transposition protocol**

This protocol delineates the steps to perform a small (20μl) "- enzyme" negative control reaction and a large (80μl) "+ enzyme" experimental mutagenesis reaction in parallel. The enzymes and buffers are mixed together in a master mix (5.5-reaction sized, to account for pipetting error) and later split into appropriate reaction sizes.

1) Make reaction mix. Combine:
   - 17.6 μl target DNA (880 ng) (*DCW1* entry vector)
   - 8.8 μl Tn7 donor DNA (220 ng) (pRZ101)
   - 11 μl 20mM ATP
   - 11 μl 20mM DTT
   - 56.1 μl ddH_2O

2) Aliquot reaction mix into two PCR tubes. Dispense 76μl into the "+ enzyme" reaction tube, and 19μl into the "- enzyme" reaction tube.

3) Make the enzyme mixture. Combine:
   - 7.49 μl TnsA
   - 3.31 μl Storage buffer A
   - 2.00 μl TnsB
   - 6μl Storage buffer B
   - 8μl $TnsC^{A255V}$
   - 8μl Storage buffer C
   - 5.2μl 50% glycerol

   Mix by flicking tube gently. Keep on ice while setting up transposition reaction.

   Note 1: The ratio of TnsA, TnsB, and $TnsC^{A255V}$ in this mixture was determined empirically. You may need to optimize the relative amount of each enzyme using your purified enzyme stocks.

   Note 2: You may refreeze the enzyme mixture at -80°C and rethaw twice, but will have decreased transposition efficiency with each thaw. We do not recommend refreezeing the stock solution of the individual Tns enzymes, so take care to store these in small aliquots so as to avoid wasting purified enzyme.

4) Make the buffer mixture for the "- enzyme" control. Combine:
   - 10.8 μl Storage buffer A
   - 8.0 μl Storage buffer B
   - 16.0 μl Storage buffer C
   - 5.2 μl 50% glycerol

5) Add 4μl enzyme mix to the "+ enzyme" tube. Flick tube to mix.
6) Add 1μl of the buffer mixture to the "- enzyme" tube. Flick tube to mix.
7) Incubate both tubes at 37°C for 10 minutes on a PCR heat block.
8) Add 300mM MgOAc to the tubes:
   For "+ enzyme" reactions, add 4.2μl 300mM MgOAc
   For "- enzyme" reaction, add 1.05μl 300mM MgOac
9) Incubate at 37°C for 1 hour on a PCR block
10) Incubate at 75°C for 5 minutes on a PCR block to heat-kill the enzymes.

Clean-up of transposition reactions

11) Transfer "+ enzyme" and "- enzyme" reactions to 1.5ml microfuge tubes. Bring the volume of each up to 100µl.
12) Add 100µl phenol:chloroform:IAA. Vortex to mix.
13) Spin 5 minutes, 4°C, 13500 rpm in a microfuge. Remove and discard organic (bottom) layer.
14) Add 100µl chloroform. Vortex to mix.
15) Spin 5 minutes, 4°C, 13500rpm in microfuge. Transfer aqueous (top) later to a new 1.5ml microfuge tube.
16) Add 10µl 3M NaOAc. Add 220µl ice-cold 100% EtOH. Chill at -20°C for 15 minutes.
17) Spin 15 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant.
18) Wash DNA pellet with 500µl ice-cold 70% EtOH.
19) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant. Allow pellet to air dry.
20) Resuspend DNA pellets in desired volume of 10mM Tris pH 7.5.
    For "+ enzyme" reactions, use 8 µl Tris pH 7.5
    For "- enzyme" reaction, use 4 µl Tris pH 7.5
21) Optional: Cleaned transposition reactions may be stored at -20°C prior to transformation.

Transformation of transposition reactions

Transform cleaned transposition reaction DNA into Invitrogen MegaX DH10B T1R electrocompetent *E. coli* cells. One set of steps is described for the "+ enzyme" reactions, another for "- enzyme" reactions. They can be performed in parallel, but are separated here for clarity.

For "- enzyme" reactions:

22A) Combine 2μl DNA with 20μl MegaX cells in a chilled electroporation cuvette.
23A) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
24A) Add 1ml Oxoid Isosensitest media to the cuvette and pipet up and down to resuspend.
25A) Transfer 900μl cells to a 1.5ml microfuge tube.
26A) Recover at 37°C for 1 hour.
27A) Plate 9μl (1% of total cells) onto three types of selective media – LB+Kan, Iso + Tmp, and Iso+Tmp+Kan .
28A) Grow plates overnight at 37°C.

For "+ enzyme" reactions:

22B) Combine 4μl DNA with 40μl MegaX cells in one chilled electroporation cuvette.
23B) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
24B) Add 1ml Oxoid Isosensitest media to the cuvette and pipet up and down to resuspend.
25B) Transfer 900μl cells to a 500ml flask containing 150 mls Isosensistest media.
26B) Recover at 37°C for 1 hour
27B) Plate 150μl (0.1% of total cells) onto three types of selective media – LB+Kan, Iso + Tmp, and Iso+Tmp+Kan
28B) Add 150μl of Kan (30mg/ml stock) and 300μl Tmp (5mg/ml stock) to the flask.
29B) Grow plates and culture overnight at 37°C.

30) The following morning, count the colonies growing on each plate. Calculate the number of independent transformants in your "+ enzyme" pools.
31) Make a glycerol frozen stock of the "+ enzyme" overnight culture. Store at -80°C.
32) Pellet the remainder of the "+ enzyme" overnight culture in a 250ml conical tube. Spin 15 minutes at 3000 rpm. Pellet can be stored at -20°C, or continue to step 33 directly.
33) Purify the DNA from the "+ enzyme" culture cells using a Qiagen HiSpeed Midiprep Kit. Elute in 1ml 10mM Tris. This is the mutagenized *DCW1**FLAG entry vector pool (ep).

Note: You can monitor the plasmid population content by restriction enzyme digestion of the DNA pool. We have observed about 50% of the plasmid pool will be unmutagenized target vector. We suspect that the MegaX may take up multiple plasmids during transformation and do not segregate them properly.

Gateway LR mobilization to create expression pools

Linearizing the mutagenized entry pool increases the efficiency of the Gateway LR reaction. Additionally, it minimizes the amount of Kan-marked plasmid that is present in the subsequent pools.

Linearize mutagenized DNA pool

1) Identify a restriction enzyme that cuts only in the backbone of the mutagenized *DCW1**FLAG pool. We chose ApaLI, which cuts in the origin of replication.
2) Digest 1μg of DNA from the *DCW1**FLAG mutagenized entry pool in a 20μl reaction. Include a negative control – use 100ng of DNA in a 20μl reaction.
3) Digest 37°C for 1 hour.
4) Run the entire negative control reaction, and 2μl of the ApaLI digestion, on a gel to verify the backbone is fully digested. If digestion was successful, proceed.

Clean the linearized DNA

5) Add 82μl ddH2O to the remaining ApaLI-digested DNA, to bring it to a final volume of 100μl.
6) Add 100μl phenol:chloroform:IAA. Vortex to mix.
7) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Remove and discard organic (bottom) layer.
8) Add 100 μl chloroform. Vortex to mix.
9) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Transfer aqueous (top) later to a new 1.5ml microfuge tube.
10) Add 10 μl 3M NaOAc. Add 220μl ice-cold 100% EtOH. Chill at -20°C for 15 minutes.
11) Spin 15 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant.
12) Wash DNA pellet with 500μl ice-cold 70% EtOH.
13) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant. Allow pellet to air dry.
14) Resuspend DNA in 20μl 10mM Tris. DNA will be 45ng/μl.

Gateway LR Reaction – Move ORF to destination vector

1 reaction volume uses 150ng destination vector and 150ng (mutagenized, linearized) entry vector

15) Make a 2.5x reaction mixture DNA, which will be split into +LR and –LR reactions. Combine: 371.25ng (8.25μl) mutagenized, linearized entry vector
    371.25ng pRZ159 (*DCW1* destination vector)
    Bring to 20μl total volume with TE
16) For the "+LR" experimental reaction, combine 16μl of the DNA mixture from step 15 and 4μl Invitrogen LR Clonase II in a PCR tube. For the "-LR" control reaction, combine 4μl DNA mixture from step 15 and 1μl TE in a PCR tube. Incubate both reactions overnight at 25°C on the PCR block.
17) Stop the LR reaction by adding proteinase K to the reactions: Add 2μl proteinase K to the "+LR" experimental reaction; Add 0.5μl proteinase K to the "-LR" negative control.
18) Incubate at 37°C for 10 minutes on the PCR block.
19) Clean reactions: Transfer reactions to 1.5ml microfuge tubes. Perform phenol/chloroform extraction and EtOH precipitation as described before (steps 5-13 under "Clean linearized DNA")
20) Resuspend "+LR" experimental DNA in 6μl TE (final [DNA] = 50ng/ul).
21) Resuspend "-LR" negative control DNA in 4μl TE (final [DNA] = 18.75ng/ul).

Transform Gateway LR reactions

Transform cleaned LR reaction DNA into Invitrogen MegaX DH10B T1R electrocompetent *E. coli* cells. One set of steps is described for the "+ enzyme" reactions, another for "- enzyme" reactions. They can be performed in parallel, but are separated here for clarity.

Note that the Car and Tmp selections (intended to select only those expression plasmids with a DCW1 orf and a Tn insertion) are performed in sequential rounds of transformation. We found simultaneous Car and Tmp drug selection seemed to pressure cells into maintaining both an unmutagenized DCW1 expression vector and a mutagenized entry vector.

For "- enzyme" reactions:

22A) Combine 2μl DNA with 20μl MegaX cells in a chilled electroporation cuvette.
23A) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
24A) Add 1ml Recovery media (provided with MegaX cells) to the cuvette and pipet up and down to resuspend.
25A) Transfer 900μl cells to a 1.5ml microfuge tube.
26A) Recover at 37°C for 1 hour.
27A) Plate 9μl (1% of total cells) onto three types of selective media – LB+Kan, Iso + Tmp, and Iso+Tmp+Kan .
28A) Grow plates overnight at 37°C.

For "+ enzyme" reactions:

22B) Combine 2μl DNA with 20μl MegaX cells in one chilled electroporation cuvette.
23B) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
24B) Add 1ml Recovery media (provided with MegaX cells) to the cuvette and pipet up and down to resuspend.
25B) Transfer 900μl cells to a 1.5ml microfuge tube.
26B) Recover at 37°C for 1 hour
27B) Transfer contents of microfuge tube to a 500ml flask containing 150mls LB. Swirl to mix.
28B) Plate 150μl (0.1% of total cells) onto five types of selective media – LB+Car, LB+Kan, Iso + Tmp, and Iso+Tmp+Kan , Iso+Tmp+Car
29B) Add 150μl of Car (100mg/ml stock) to the flask.
30B) Grow plates and culture overnight at 37°C.

30) The following morning, count the colonies growing on each plate. Calculate the number of independent transformants in the "+ enzyme" pools.
31) Make a glycerol frozen stock of the "+ enzyme" overnight culture. Store at -80°C.
32) Pellet the remainder of the "+ enzyme" overnight culture in a 250ml conical tube. Spin 15 minutes at 3000 rpm. Pellet can be stored at -20°C, or continue to step 33 directly.
33) Purify the DNA from the "+ enzyme" culture cells using a Qiagen HiSpeed Midiprep Kit. Elute in 1ml 10mM Tris. This is the expression pool, selected for Car^R ("xpC").

Retransform for 2nd round drug selection: Select expression pool for plasmids with Tn insertions

34) Combine 1μl "xpC" expression pool DNA (50ng/μl) with 20μl MegaX cells in one chilled cuvette.
35) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
36) Add 1ml Oxoid Isosensitest media to the cuvette and pipet up and down to resuspend.
37) Transfer 900μl cells to 150mls Isosensitest media (in a 500ml flask).
38) Recover at 37°C for 1 hour
39) Plate 150μl (0.1% of total cells) onto five types of selective media – LB+Car, LB+Kan, Iso + Tmp, and Iso+Tmp+Kan , Iso+Tmp+Car
40) Make a 1:1000 dilution of recovered cells (2μl cells into 2ml Isosensitest media). Plate 150μl of a 1:1000 dilution of cells onto Iso+Tmp, LB+Car, and Iso+Tmp+Car plates
41) Add 300μl of Tmp (5mg/ml stock) to the flask.
42) Grow plates and culture overnight at 37°C.
43) The following morning count the colonies growing on each plate. Calculate the number of independent transformants in your pools.
44) Make a glycerol frozen stock of the overnight culture. Store at -80°C.
45) Pellet the remainder of the overnight culture in a 250ml conical tube. Spin 15 minutes at 3000 rpm. Pellet can be stored at -20°C, or continue to step 46 directly.
46) Purify the DNA from the pelleted cells using a Qiagen HiSpeed Midiprep Kit. Elute in 1ml 10mM Tris. This is the expression pool, now selected for Tmp^R ("xpT").

Restriction digestions to remove Tn7 DNA from plasmids

At this point, the pools contain expression plasmids with mutagenized target ORF. This is based on the ability to grown in MegaX cells – confirming the expression backbone instead of the entry or destination vectors, and the selection for both Car^R and Tmp^R - confirming the expression backbone (Car^R) and the presence of Tn7 (Tmp^R). At this time, the Tn7 ends and Tmp^R gene are no longer needed, and are removed from the plasmids using a series of restriction digests.

Remove Tn7L end with *Fse*I digest

1) Digest 500ng of xpT pool with *Fse*I; use 100ng of xpT pool for a "no enzyme" control.

No enzyme	***Fse*I**	
100 ng	500 ng	xpT pool DNA
1 μl	1 μl	10x NEB buffer #4
1 μl	1 μl	10x BSA
0 μl	0.5 μl	*Fse*I
To 10μl total	To 10μl total	dH2O

2) Incubate restriction digests and control reactions at 37°C for 1hr.
3) Heat inactivate *Fse*I by incubating reactions at 65°C for 20 minutes.
4) Run all of the "no enzyme" control and 2μl (100ng of DNA) from *Fse*I digest on an agarose gel to confirm restriction digest was successful. The uncut control should contain supercoiled plasmid, and the *Fse*I digest should have a high MW band of linearized plasmid DNA; the Tn7L fragment is 211bp long, though this fragment is sometimes not visible. If successful, continue with protocol.

Ligate to recircularize plasmid after *Fse*I digestion

5) Add 12μl of dH2O to remaining *Fse*I-digested xpT material to bring final volume to 20μl.
6) Set up ligation reaction, and a negative control, as shown:

- control	**+ ligase**	
10 μl	10 μl	DNA (200ng)
5 μl	5 μl	10x T4 ligase buffer
5 μl	5 μl	10mM ATP
0 μl	2.5 μl	T4 DNA ligase
30 μl	27.5 μl	dH2O
50 μl	50 μl	TOTAL

7) Ligate at room temperature, 30 minutes.
8) Heat inactivate ligase enzyme by incubating at 65°C, 20 minutes

Clean the ligated DNA after removal of Tn7L

9) Add 50μl ddH2O to the ligated DNA, to bring it to a final volume of 100μl.
10) Add 10 μl 3M NaOAc. Add 220μl ice-cold 100% EtOH. Chill at -20°C for 15 minutes.
11) Spin 15 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant.
12) Wash DNA pellet with 500μl ice-cold 70% EtOH.
13) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant. Allow pellet to air dry.
14) Resuspend DNA in 4μl 10mM Tris. DNA will be roughly 50ng/μl

Transform ligated DNA after removal of Tn7L

For both "- ligase controls" and "+ ligase" reactions:

15) Combine 1μl of DNA (at 50ng/μl) with 20μl MegaX cells in a chilled electroporation cuvette.
16) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
17) Add 1ml Isosensitest to the cuvette and pipet up and down to resuspend.
18) Transfer 900μl cells to a 1.5ml microfuge tube.
19) Recover at 37°C for 1 hour.
20) Make a 1:10 dilution of recovered cells (22μl cells + 200μl Isosensitest). Plate 90μl (1% of total cells) onto prewarmed LB+Kan and Iso+Kan+Tmp plates

For "-ligase" control reactions:

21A) Make a 1:1000 dilution of recovered cells (3μl of the 1:10 diluted cells + 297μl Isosensitest). Plate 90μl (0.01% of total cells) onto prewarmed LB+Car, Iso+Tmp, and Iso+Tmp+Car plates.

For "+ligase" reactions:

21B) Make a 1:10,000 dilution of recovered cells (1μl of the 1:10 diluted cells + 999μl Isosensitest). Plate 90μl (0.001% of total cells) onto prewarmed LB+Car, Iso+Tmp, and Iso+Tmp+Car plates.

22) Transfer remainder of the "+ligase" reaction to a flask containing 150ml Isosensitest. Add 300μl 500x Tmp stock.
23) Grow plates and culture at 37°C overnight.
24) The following morning, pull plates and count colonies. Calculate the number of independent transformants in your "+ ligase" pools.
25) Make a glycerol frozen stock of the "+ ligase" overnight culture. Store at -80°C.
26) Pellet the remainder of the "+ ligase" overnight culture in a 250ml conical tube. Spin 15 minutes at 3000 rpm. Pellet can be stored at -20°C, or continue to step 37 directly.
27) Purify the DNA from the "+ ligase" culture cells using a Qiagen HiSpeed Midiprep Kit. Elute in 1ml 10mM Tris. This is the expression pool after removal of Tn7L ("xpT-L").

Remove Tn7R end with *Pme*I digest

21) Digest 1 μg of xpT-L pool with *Pme*I; use 100ng of xpT-L pool for a "no enzyme" control.

No enzyme	***Pme*I**	
100 ng	1000 ng	xpT pool DNA
1 μl	1 μl	10x NEB buffer #4
1 μl	1 μl	10x BSA
0 μl	0.5 μl	*Pme*I
To 10μl total	To 10μl total	dH2O

22) Incubate restriction digests and control reactions at 37°C for 1hr.
23) Heat inactivate *Pme*I by incubating reactions at 65°C for 20 minutes.
24) Run all of the "no enzyme" control and 1μl (100ng of DNA) from *Pme*I digest on an agarose gel to confirm restriction digest was successful. The uncut control should contain supercoiled plasmid. The *Pme*I digest should have a high MW band of plasmid backbone DNA, and the released Tn7R end should be visible at 863bp. If successful, continue with protocol.

Clean digest prior to ligation

We found that residual PmeI may interfere with the success of the ligation reaction.

25) Add ddH2O to remaining *Pme*I-digested xpT-L material to bring final volume to 100μl.
26) Add 100μl phenol:chloroform:IAA. Vortex to mix.
27) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Remove and discard organic (bottom) layer.
28) Add 100 μl chloroform. Vortex to mix.
29) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Transfer aqueous (top) later to a new 1.5ml microfuge tube.
30) Add 10 μl 3M NaOAc. Add 220μl ice-cold 100% EtOH. Chill at -20°C for 15 minutes.
31) Spin 15 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant.
32) Wash DNA pellet with 500μl ice-cold 70% EtOH.
33) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant. Allow pellet to air dry.
34) Resuspend DNA pellet (may not be visible) in 20μl of 10mM Tris pH 7.5.

Ligate to recircularize plasmid after *Pme*I digestion

Longer overnight reaction is performed to encourage ligation of blunt ends from PmeI digest.

35) Set up ligation reaction (and a negative control) as shown:

- control	**+ ligase**	
10 μl	10 μl	DNA (200ng)
5 μl	5 μl	10x T4 ligase buffer
5 μl	5 μl	10mM ATP
0 μl	2.5 μl	T4 DNA ligase
30 μl	27.5 μl	dH2O
50 μl	50 μl	TOTAL

36) Ligate at 16°C, 16 hours.
37) Heat inactivate ligase enzyme by incubating at 65°C, 20 minutes

Clean the ligated DNA after removal of Tn7R

38) Add 50μl ddH2O to the ligated DNA, to bring it to a final volume of 100μl.
39) Add 10 μl 3M NaOAc. Add 220μl ice-cold 100% EtOH. Chill at -20°C for 15 minutes.
40) Spin 15 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant.
41) Wash DNA pellet with 500μl ice-cold 70% EtOH.
42) Spin 5 minutes, 4°C, 13500 rpm in microfuge. Discard supernatant. Allow pellet to air dry.
43) Resuspend DNA in 4μl 10mM Tris. DNA will be roughly 50ng/μl

Transform ligated DNA after removal of Tn7R

For both "- control" and "+ ligase" reactions:

44) Combine 1μl of DNA (~ 112ng/μl) with 20μl MegaX cells in a chilled electroporation cuvette.
45) Electroporate at 2.0 kV, 200Ω, 25μF, according to protocol.
46) Add 1ml Isosensitest to the cuvette and pipet up and down to resuspend.
47) Transfer 900μl cells to a 1.5ml microfuge tube.

For "-ligase" control reactions:

48A) Recover at 37°C for 1 hour in microfuge tube.
49A) Make 1:10 dilution of recovered cells (30μl cells + 270μl LB). Plate 90μl (1% of total cells) onto prewarmed LB+Kan and Iso+Tmp+Kan plates.
50A) Make a 1:100 dilution of recovered cells (40μl of 1:10 dilution of cells, 360μl LB). Plate 90μl (0.1% of total cells) onto prewarmed LB+Car, Iso+Tmp, and Iso+Tmp+Car plates
51A) Make a 1:1000 dilution of recovered cells (40μl of 1:100 dilution of cells, 360μl LB). Plate 90μl (0.01% of total cells) onto prewarmed LB+Car, Iso+Tmp, and Iso+Tmp+Car plates.
52A) Grow plates at 37°C overnight

For "+ligase" control reactions:

48B) Transfer the cells to a 500ml flask containing 150mls LB. Recover at 37°C for 1 hour.
49B) Plate 150μl of recovered cells (0.1% of total cells) onto prewarmed LB+Kan and Iso+Tmp+Kan plates.
50B) Make a 1:10 dilution of recovered cells (60μl of cells, 540μl LB). Plate 150μl (0.01% of total cells) onto prewarmed LB+Car, Iso+Tmp, and Iso+Tmp+Car plates
51B) Make a 1:100 dilution of recovered cells (50μl of 1:10 dilution of cells, 450μl LB). Plate 150μl (0.001% of total cells) onto prewarmed LB+Car, Iso+Tmp, and Iso+Tmp+Car plates.
52B) Add 150μl 100x Car stock to culture flask. Grow plates and flask at 37°C overnight

53) The following morning, pull plates and count colonies. Calculate the number of independent transformants in your "+ ligase" pools.
54) Make a glycerol frozen stock of the "+ ligase" overnight culture. Store at -80°C.
55) Pellet the remainder of the "+ ligase" overnight culture in a 250ml conical tube. Spin 15 minutes at 3000 rpm. Pellet can be stored at -20°C, or continue to step 37 directly.
56) Purify the DNA from the "+ ligase" culture cells using a Qiagen HiSpeed Midiprep Kit. Elute in 1ml 10mM Tris. This is the expression pool after removal of both Tn7L and Tn7R("xpT-L-R").

Diagnostic digests

We highly recommend performing diagnostic restriction digests on plasmid pools from each phase of this protocol. This helps to characterize the proportion of each pool that is mutagenized, identify any lingering donor plasmid that persists in the expression pools, and helps determine the complexity of the pool (ie, is the Tn inserted at many locations throughout the target ORF?).

We suggest performing many restriction digests in parallel. Use an enzyme that cuts only in the donor backbone, one that cuts only in the expression backbone, one that cuts in the Tn7L, and one that cuts in Tn7R. Perform "no enzyme" controls, as well. Include one double digestion which cuts in the epitope tag and one in the backbone – this will determine the complexity of the mutagenized pool. Perform restriction digest on the following DNA samples:

1) Unmutagenized target donor vector
2) wt expression vector
3) mutagenized donor pool
4) xpC pool
5) xpT pool
6) xpT-L pool
7) xpT-L-R pool

References

1. Brachmann, C.B. et al. 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**, 115-132 (1998).
2. Winzeler, E.A. et al. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science* **285**, 901-906 (1999).
3. Gamas, P. & Craig, N.L. Purification and characterization of TnsC, a Tn7 transposition protein that binds ATP and DNA. *Nucleic acids research* **20**, 2525-2532 (1992).
4. Choi, K.Y., Li, Y., Sarnovsky, R. & Craig, N.L. Direct interaction between the TnsA and TnsB subunits controls the heteromeric Tn7 transposase. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E2038-2045 (2013).

File S2

Calculating fold coverage of Tn-mutagenized pools.

For entry pools: Fold coverage was calculated by dividing the number of Tmp^RKan^R transformants by the length of plasmid DNA "available" for mutagenesis. The target plasmid, *DCW1* entry vector, is 3608bp long. When considering the portions of the entry vector into which we could expect insertions, we excluded the origin (683 bp) and Kan^R gene (807bp), because plasmids with insertions in these regions would not be expected to propagate. Thus, 4236bp (2118 bp x 2 strands of DNA) of sequence in *DCW1* entry vector was available for mutagenesis. We observed 2.01×10^5 Tmp^RKan^R colonies / 4236bp = 47-fold coverage.

For expression pools: There are 1373 bp between Gateway recombination sites in a wild type *DCW1* expression vector. In our mutagenized expression pools, all Tn7 insertions will be present in this area, since that is the only region mobilized from the mutagenized *DCW1* entry pool by Gateway recombination. Because the Tn7 can insert in either direction, we consider the "available" DNA for mutagenesis in these constructs to be 2746 bp long. Fold coverage is calculated using the number of Tmp^RCar^R colonies in the expression pools, or Car^R colonies in the final pool. See Supplementary Note 3 for a discussion of how the number of Tmp^RCar^R colonies was calculated in the xpC pool. We observed 91x fold coverage in the initial expression pool (xpC), and recovered sufficiently numbers of transformants in subsequent steps to maintain this coverage. As the initial mutagenesis of the *DCW1* entry vector had 47x coverage, these expression pools are sufficiently large to maintain the complexity present in the initial mutagenized entry pool.

File S3

Revising xpC pool sizes

To properly assess the size of the xpC pool, we found it was necessary to plate on selective media both pre-and post- Car selection. We suspect the MegaX *E. coli* cells are transformed with multiple plasmids and do not segregate them properly; thus the need to distinguish between cells that carry Tmp^R and Car^R on a single plasmid (as desired), and those that carry it on separate plasmids. Because of the possibility the strain carries multiple plasmids, co-selection with Tmp and Car may overestimate the true number of mutagenized expression plasmids in the pool (as counted from Iso+Tmp+Car plates). Theoretically, by performing the Car and Tmp selection sequentially, we minimize the selective pressure for MegaX cells to maintain multiple plasmids.

First, to assess the transformation efficiency, recovered cells were plated (prior to Car selection overnight) and found to contain $5x10^6$ Car^R colonies. This represents the number of independent expression vectors (mutagenized or not) in the xpC pool. Second, to determine how many of these Car^R colonies are also Tmp^R, we took a portion of the cells from the xpC pool (post-Car selection) onto LB+Car plates, and then replica plated to various selective media to determine the precise drug-resistance of each colony (Table S5).

In detail: cells from the xpC frozen stock were resuspended in PBS, diluted, and plated to get roughly 200 colonies per LB+Car plate; 4 LB+Car plates were used. These plates were grown overnight at 37°C. Each LB+Car plate was then replica plated (in order) to LB+Kan, Iso+Tmp+Kan, Iso+Tmp, Iso+Tmp+Car, and LB+Car plates, and grown at 37°C overnight. The total number of colonies were counted on each plate. Additionally the Iso+Tmp+Car and Iso+Tmp+Kan were compared to identify any colonies resistant to Tmp, Car, and Kan. This would suggest a single cell contains two plasmid populations, indicating sub-optimal plasmid segregation in MegaX cells.

Calculations for drug resistance of colonies were performed as follows:

$Tmp^RKan^RCar^R$ colonies = # colonies on Iso+Tmp+Kan plate (is Car^R, based on growth on source plate).

Tmp^RKan^R colonies = (colonies on Iso+Tmp+Kan plates) – (# of $Tmp^RKan^RCar^R$ colonies)

Tmp^RCar^R colonies = (colonies on Iso+Tmp+Car plates) –(# of $Tmp^RKan^RCar^R$ colonies)

Kan^RCar^R colonies = (# colonies on LB+Kan plate) – (# of $Tmp^RKan^RCar^R$ colonies) – (# of Tmp^RKan^R colonies)

Car^R only colonies = (# colonies on LB+Car plate) – (# of $Tmp^RKan^RCar^R$ colonies) – (# of Tmp^RCar^R colonies)-(# of Kan^RCar^R colonies).

See results of calculations in Table S5. The correction factor was calculated as the Tmp^RCar^R/ (total Car^R), as we selected for only Car^R in the culture and we can assume we counted the LB+Car plates accurately. "total Car^R" would be the sum of ($Tmp^RKan^RCar^R$), (Tmp^RCar^R), (Kan^RCar^R) and (Car^R only) colonies. We found that an average 5% of the Car^R colonies in the xpC culture were actually Tmp^RCar^R. Thus, 5% of the $5x10^6$ Car^R transformants counted would estimate our Tmp^RCar^R population to be $2.5x10^5$ Tmp^RCar^R colonies in the xpC pool, as shown in Table 1.

File S4

Assessing pool complexity

Restriction digests were used diagnostically at each step of the mutagenesis and processing of the mutagenized pool. This enabled us to monitor if donor vectors or unmutagenized expression plasmids were errantly maintained in the pool. Also, double digests with enzymes that cut in the backbone and in the epitope tag allowed us to qualitatively assess pool complexity. Digesting a mutagenized plasmid pool in this way should result in smeared bands after gel electrophoresis, where the range of fragment sizes should reflect epitope tags inserted throughout the target ORF (data not shown).