Short Technical Reports

Vector for Pop-In/ Pop-Out Gene Replacement in *Pichia pastoris*

BioTechniques 31:306-312 (August 2001)

ABSTRACT

Gene replacement in yeast is often accomplished by using a counterselectable marker such as URA3. Although *ura3* strains of *Pichia pastoris* have been generated, these strains are inconvenient to work with because they grow slowly, even in the presence of uracil. To overcome this limitation, we have developed an alternative counterselectable marker that can be used in any *P. pastoris* strain. This marker is the *T-urf13* gene from the mitochondrial genome of male-sterile maize. Previous work showed that expression of a mitochondrially targeted form of *T-urf13* in *Saccharomyces cerevisiae* rendered the cells sensitive to the insecticide methomyl, and similar results have now been obtained with *P. pastoris*. We have incorporated *T-urf13* into a vector that also contains an *ARG4* marker for positive selection. The resulting plasmid allows for pop-in/pop-out gene replacement in *P. pastoris*.

INTRODUCTION

Because yeast cells are proficient at homologous recombination, simple transformation protocols can be used to replace the chromosomal copy of a gene with a modified version of the gene (12). This technique allows researchers to add tags or introduce mutations while expressing a gene in its normal context. A common strategy for gene replacement in *Saccharomyces cerevisiae* is the "pop-in/pop-out" method, which exploits the fact that cells expressing a wild-type *URA3* gene are unable to grow in the presence of 5-fluoroorotic acid (5-FOA) (1). Gene replacement is performed by transforming a linearized plasmid into a *ura3* auxotroph; positive selection for the *URA3* marker yields clones that have integrated the plasmid, and subsequent counterselection in the presence of 5-FOA yields clones that have excised the plasmid and restored the *ura3*

phenotype. Thus, the pop-in/pop-out method has the advantage of regenerating the selectable marker. The *URA3*/5-FOA system can potentially be used for gene replacement in budding yeasts other than *S. cerevisiae*. One example is *Pichia pastoris*, which has become a popular model organism for protein production and cell biological studies (5). Unfortunately, however, *ura3* auxotrophs of *P. pastoris* grow slowly, even in the presence of uracil, making these strains cumbersome to use (7). We therefore sought an alternative counterselectable marker that would function in any *P. pastoris* strain.

A promising candidate was the maize mitochondrial gene *T-urf13*, which encodes a 13-kDa polypeptide that confers T-type cytoplasmic male sterility (9). Maize strains expressing *T-urf13* are sensitive to certain fungal toxins and to the carbamate insecticide methomyl. Previous investigators recapitulated these phenotypes in *S. cerevisiae* by linking a mitochondrial targeting presequence to the *T-urf13* gene product (3,6). Expression of *T-urf13*

had little effect on *S. cerevisiae* cells under normal growth conditions, but addition of methomyl uncoupled the mitochondria, thereby preventing respiratory growth on carbon sources such as glycerol. We now report similar results with *P. pastoris*. This conditional lethality allows *T-urf13* to be used as a counterselectable marker. By incorporating *T-urf13* into a vector that also contains a positive selectable marker, we have achieved pop-in/pop-out gene replacement in *P. pastoris*.

MATERIALS AND METHODS

Yeast Strains and Growth Conditions

All experiments were performed with the haploid *P. pastoris* strain PPY12, which has the genotype *his4 arg4* (4). PPY12 cells were grown in liquid YPG medium (1% yeast extract, 2% peptone, 3% glycerol, 20 mg/L adenine sulfate, 20 mg/L uracil), then transformed by electroporation (13)

Figure 1. Map of the plasmid pPOP. This plasmid is a derivative of pUC19. The indicated restriction sites are unique except for *SacI* and *HincII*.

and plated on selective SG medium lacking arginine [0.67% yeast nitrogen base, 3% glycerol, 1×CSM-Arg (complete supplement mixture lacking arginine; Qbiogene, Carlsbad, CA, USA), 2% agar]. Where indicated, methomyl was added to SG medium to a final concentration of 1 mM immediately before the plates were poured. Because methomyl is heat sensitive, it was added after the medium had cooled to about 50°C. Purified methomyl was purchased from Sigma-Aldrich (St. Louis, MO, USA). Essentially identical results were obtained with the liquid pesticide Lannate LV, which contains methomyl as its active component; our Lannate sample (a kind gift of DuPont Agricultural Products, Wilmington, DE, USA), which had a methomyl content of 29% (1.8 M), was used for all of the experiments described in this paper. As soon as methomyl-containing plates had solidified, they were stored in the dark at 4°C. These plates were used within a week of being poured.

Plasmids

A cassette encoding the mitochondrial targeting presequence of *Neurospora crassa* ATPase subunit 9 fused to *T-urf13* was amplified by PCR from

plasmid pEMBLYe30/2-21urf13-TW (3) using *Pfu* DNA polymerase (Stratagene, La Jolla, CA, USA) and the oligonucleotides 5'-ACAAAATGGC-CTCCACTCGTGTCTCG-3' (sense) and 5'-CTATCACGGTACTTGTACGCTATCGGTC-3' (antisense). This PCR product was inserted into the integrating vector pIB2 (13) that had been digested with *Eco*RI and *Hind*III and blunted with Klenow enzyme. The resulting construct, designated pIB2-T-urf13, contains the chimeric *T-urf13* gene under control of the strong constitutive *GAP* promoter and the *AOX1* transcription terminator. This plasmid was linearized with *Stu*I for integration into the *his4* locus of PPY12 cells.

pPOP was constructed as follows. A cassette containing the chimeric *T-urf13* gene flanked by the *GAP* promoter and the *AOX1* transcription terminator was amplified from pIB2-T-urf13 using *Pfu* DNA polymerase and the oligonucleotides 5'-GATCCT-TTTTGTAGAAATGTCTTGGTG-3' (sense) and 5'-ACATGTGTGGGAAATACCAAGAAAAAC-3' (antisense). This PCR product was inserted into pUC19 (New England Biolabs, Beverly, MA, USA) that had been digested with *Hinc*II. The resulting construct was digested with *Xba*I plus *Pst*I, and

Figure 2. Methomyl is lethal to *P. pastoris* cells that are expressing T-urf13. *P. pastoris* PPY12 cells were transformed with pIB2-T-urf13 or with the parental pIB2 vector as a control. Representative clones that had undergone correct integration into the *his4* locus (13) were spotted by 3-fold serial dilutions onto SG plates lacking histidine and either lacking or containing 1 mM methomyl. The clone expressing T-urf13 failed to grow in the presence of methomyl. In the absence of methomyl, cells expressing T-urf13 grew almost as rapidly as the control cells.

Short Technical Reports

the excised 1393-bp fragment was blunted with T4 DNA polymerase and inserted into pSG464 (4) that had been cut with *Nde*I plus *Sna*BI and blunted with Klenow enzyme. The sequence of pPOP has been deposited in GenBank® under the accession no. AF323603.

To replace *SEC7* with *SEC7-GFP*, pPOP was first mutagenized using the QuikChange™ kit (Stratagene) to eliminate the *Xmn*I site. An *Eco*RI-*Hind*III fragment spanning the 3' end of *P. pastoris SEC7* (GenBank accession no. AF356651) was then inserted into the corresponding sites in the mutagenized pPOP vector. As described elsewhere (10), a triple-*GFP* cassette was inserted at the end of the *SEC7* gene. The resulting construct was linearized with *Xmn*I for integration into the *SEC7* locus. To replace *SEC12* with *SEC12-myc*, a *Bam*HI fragment containing the 3' portion of *P. pastoris SEC12* fused to a c-myc epitope sequence (11) was inserted into the *Bam*HI site of pPOP. The resulting construct was linearized with *Xho*I for integration into the *SEC12* locus.

Pop-In/Pop-Out Gene Replacements

For the *SEC7* and *SEC12* gene replacements, PPY12 cells were transformed with the appropriate linearized vector and selected for growth on SG-Arg. A typical transformation of this type using 0.5–1 µg linearized plasmid DNA yielded 25–100 transformants. Some of these transformants represented undesired events, presumably including gene conversion at the *ARG4* locus or random insertion of the plasmid into genomic DNA, but 25%–50% of the transformants were bona fide pop-in clones. To identify the desired clones, five transformants were tested for growth on SG-Arg plates containing 1 mM methomyl, and the methomyl-sensitive clones were examined by fluorescence microscopy for the expected localization pattern (11). Pop-in clones were then plated as follows to select for the desired pop-out clones. A small clump of cells was suspended in 200 µL sterile water and plated onto SG medium containing arginine plus 1 mM methomyl to select for cells that had excised the *T-urf13* gene. Viable clones were examined by fluorescence microscopy to confirm that the tagged

version of the gene was still present and were tested to confirm that they could no longer grow on SG-Arg, indicating that the *ARG4* gene from the plasmid had been excised.

To verify that the expected chromosomal rearrangements had occurred during the replacement of *SEC7* with *SEC7-GFP*, we performed Southern analysis as follows. Yeast genomic DNA was prepared from putative pop-in and pop-out clones using the Easy-DNA™ kit (Invitrogen, Carlsbad, CA, USA). The DNA samples were digested with *Hind*III, electrophoresed on a 0.5% agarose gel, and transferred to a positively charged nylon membrane. A probe corresponding to the 3'-untranslated region of the *P. pastoris SEC7* locus was generated by PCR using the

oligonucleotides 5'-GTAGACATTGC-AGTAAGTACTAAGAC-3' (sense) and 5'-GACTTGACAGGCTTGGTAGAGACTTG-3' (antisense), and the nylon membrane was probed using the DIG system (Roche Molecular Biochemicals, Indianapolis, IN, USA) according to the manufacturer's instructions. A similar strategy was used to verify the chromosomal arrangements during the replacement of *SEC12* with *SEC12-myc*, except that the genomic DNA samples were digested with *Eco*RI plus *Avr*II, and a probe corresponding to the 3' portion of the *SEC12* coding sequence (8) was generated by PCR using the oligonucleotides 5'-CCTGATGAAATTCCTCGGAAGTTC-3' (sense) and 5'-GTCTTTTACTTCCTCAAGCACCTC-3' (antisense).

Figure 3. Schematic representation of gene replacement using pPOP. The gene of interest or a fragment of this gene is subcloned into pPOP and modified to introduce the desired mutation (represented by the small black rectangle). This construct must be linearized using a restriction enzyme that cuts uniquely within the gene, leaving a minimum of 100 bp of unaltered genomic sequence on each side of the restriction site. The linearized DNA is transformed into an *arg4* strain of *P. pastoris*. Positive selection on glycerol plates lacking arginine yields pop-in clones that have integrated the plasmid at the chromosomal locus of the gene. These pop-in clones are sensitive to methomyl. For subsequent counterselection, cells from a pop-in clone are spread on a glycerol plate containing both arginine and methomyl. Recombination between the duplicated sequences excises the pPOP vector, yielding *arg4* pop-out clones that are resistant to methomyl. A subset of these pop-out clones will retain the mutated version of the gene.

Short Technical Reports

RESULTS AND DISCUSSION

We have designed a vector that allows for pop-in/pop-out gene replacement in *P. pastoris*. This vector, termed pPOP, contains the *S. cerevisiae* *ARG4* gene for positive selection (2) and the *T-urf13* gene for counterselection (Figure 1). Strong constitutive expression of *T-urf13* is driven by the *GAP* promoter and *AOX1* transcription terminator from *P. pastoris* (13,15). As anticipated from previous studies of *S. cerevisiae* (3,6), 1 mM methomyl has only a slight effect on the respiratory growth of wild-type *P. pastoris* cells but completely prevents the growth of cells expressing *T-urf13* (Figure 2).

Our data confirm that the *GAP/T-urf13/AOX1* gene cassette provides a convenient counterselectable marker. This gene cassette is small (1377 bp). In the absence of methomyl, *T-urf13* has only a mild effect on growth of *P. pastoris*: during growth in liquid SG

medium at 30°C, wild-type cells have a doubling time of 2.3 h, whereas cells expressing *T-urf13* have a doubling time of 2.5 h (not shown). Counterselection with methomyl is technically straightforward and inexpensive because methomyl can be obtained in bulk as the insecticide Lannate. However, two technical issues must be considered when performing counterselection with *T-urf13*. First, methomyl is light sensitive and breaks down rapidly in aqueous solution, so methomyl-containing plates should be used as soon as possible after preparation, and any unused plates must be transferred quickly to storage in the dark at 4°C. Second, expression of *T-urf13* is mildly toxic to *P. pastoris* cells, even in the absence of methomyl. This toxicity is less pronounced in glycerol medium than in glucose medium, perhaps because glycerol yields somewhat weaker expression from the *GAP* promoter (13, 15) or perhaps because glycerol-grown

cells have more mitochondria and a correspondingly lower intramitochondrial concentration of the *T-urf13* gene product. In any case, this mild toxicity means that strains expressing *T-urf13* should not be cultured for prolonged periods but should be used immediately for counterselection.

The strategy for using pPOP to replace genes in *P. pastoris* is diagrammed in Figure 3. To confirm that pPOP works as predicted, we replaced the *SEC7* gene with a *SEC7-GFP* fusion gene. *Sec7p* is an abundant protein of the late Golgi, and our previous studies of *S. cerevisiae* revealed that *Sec7p-GFP* functions normally and gives a bright fluorescence signal (10,14). A DNA fragment spanning the 3' end of *P. pastoris* *SEC7* was subcloned into pPOP and modified by inserting a triple-*GFP* cassette at the end of the coding sequence. The resulting construct was linearized at a unique restriction site upstream of the stop codon and transformed into strain PPY12, which is an *arg4* auxotroph. Five *Arg*⁺ transformants were tested as follows to identify a correct pop-in clone. First, the transformants were tested for methomyl sensitivity by plating on glycerol medium lacking arginine but containing methomyl. All of the transformants were unable to grow in the presence of methomyl, indicating that they were expressing *T-urf13*. As judged by fluorescence microscopy, two of the five transformants showed the expected fluorescence pattern for Golgi-localized *Sec7p-GFP* (11). A transformant that had passed both of these tests was subjected to Southern analysis to verify that plasmid integration had caused the expected rearrangement at the *SEC7* locus (Figure 4). This pop-in clone was then used to generate a pop-out strain by plating on glycerol medium containing both arginine and methomyl, thereby selecting for excision of the *T-urf13* gene. Five clones that grew under these conditions were tested as follows to identify a correct pop-out strain. All five of the clones were no longer able to grow in the absence of arginine, confirming that they had also excised the *ARG4* gene. One of the five clones still showed fluorescent Golgi structures, indicating that recombination had occurred downstream

Figure 4. Southern analysis of pop-in and pop-out strains expressing *SEC7-GFP*. As described in Materials and Methods, a pPOP derivative was used to replace the *P. pastoris* *SEC7* gene with a chimeric gene in which *SEC7* was fused to three tandem copies of *GFP*. Genomic DNA samples prepared from the parental strain (PPY12), a pop-in clone, and the final pop-out strain were digested with *Hind*III and subjected to Southern analysis. (A) With the parental strain (lane 1), the probe recognizes a single 3.5-kb *Hind*III fragment that spans the 3' end of *SEC7*. With the pop-in clone (lane 2), the corresponding fragment is 5.6 kb because of addition of the triple-*GFP* cassette. The pop-in clone also yields a hybridizing fragment of 7.4 kb that includes the entire pPOP vector plus a duplicated sequence spanning the 3' end of *SEC7*. With the pop-out strain (lane 3), the probe recognizes only the 5.6-kb fragment that spans the 3' end of *SEC7-GFP*. (Because lane 3 contained less DNA than the other two lanes, the signal from lane 3 was enhanced slightly by processing the digital image.) These band patterns confirm that gene replacement occurred as diagrammed in Figure 3. (B) Schematic illustration of the genomic architecture at the *SEC7* locus in the parental, pop-in, and pop-out strains. The black rectangles represent *SEC7* sequence, the gray rectangles represent the triple-*GFP* cassette, and the white rectangle represents the pPOP vector. The arrows mark *Hind*III sites, and the small black bars indicate the sequences recognized by the probe.

Short Technical Reports

of the stop codon, thereby retaining the triple-*GFP* cassette in the chromosome. This pop-out event was verified by Southern analysis (Figure 4).

We also performed a second gene replacement using pPOP. In an earlier study, the *P. pastoris* *SEC12* gene was replaced with a *SEC12-myc* fusion gene by a pop-in strategy (11). With that approach, a duplicated fragment spanning the 3' end of *SEC12* remained in the chromosome. As described above for *SEC7*, we have now cleanly replaced the *SEC12* gene with *SEC12-myc* by the pop-in/pop-out method (not shown; see Materials and Methods). The combined data confirm that pPOP can be used for gene replacement in *P. pastoris*.

When creating a gene replacement construct based on pPOP, one should bear in mind that the pop-in and pop-out events will be efficient only if the construct is designed carefully (12). For the pop-in step, at least 100 bp of unaltered genomic sequence should be present on either side of the restriction site that is used for linearization. (To ensure that this restriction site is unique

in the final construct, any additional sites can be removed from the plasmid by directed mutagenesis. This strategy is illustrated by our replacement of *SEC7* with *SEC7-GFP*: we removed the *XmnI* site from pPOP to take advantage of the *XmnI* site near the 3' end of *SEC7*.) For the pop-out step, an important consideration is that excision of the pPOP vector sequence can yield either the desired gene replacement or a re-generated copy of the original gene, depending on the site of recombination. The construct should therefore be designed to bias the recombination event in favor of gene replacement. In the examples described above, the *GFP* and *myc* tags were inserted at the 3' ends of the genes, so each pPOP construct included a long stretch of 3'-untranslated genomic sequence to promote recombination downstream of the tag sequence. This type of pop-in/pop-out strategy takes full advantage of the efficient recombination machinery in yeast.

REFERENCES

1. Boeke, J.D., J. Trueheart, G. Natsoulis, and G.R. Fink. 1987. 5-fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* 154:164-175.
2. Cregg, J.M. and K.R. Madden. 1987. Development of yeast transformation systems and construction of methanol-utilization-defective mutants of *Pichia pastoris* by gene disruption, p. 1-18. In G.G. Stewart, I. Russell, R.D. Klein, and R.R. Hiebsch (Eds.), *Biological Research on Industrial Yeasts*. CRC Press, Boca Raton, FL.
3. Glab, N., R.P. Wise, D.R. Pring, C. Jacq, and P. Slonimski. 1990. Expression in *Saccharomyces cerevisiae* of a gene associated with cytoplasmic male sterility from maize: respiratory dysfunction and uncoupling of yeast mitochondria. *Mol. Gen. Genet.* 223:24-32.
4. Gould, S.J., D. McCollum, A.P. Spong, J.A. Heyman, and S. Subramani. 1992. Development of the yeast *Pichia pastoris* as a model organism for a genetic and molecular analysis of peroxisome assembly. *Yeast* 8:613-628.
5. Higgins, D.R. and J.M. Cregg (Eds.). 1998. *Pichia* Protocols. Humana Press, Totowa, NJ.
6. Huang, J., S.-H. Lee, C. Lin, R. Medici, E. Hack, and A.M. Myers. 1990. Expression in yeast of the T-URF13 protein from Texas male-sterile maize mitochondria confers sensitivity to methomyl and to Texas-cytoplasm-specific fungal toxins. *EMBO J.* 9:339-347.
7. Lin Cereghino, G.P., J. Lin Cereghino, A. Jay Sunga, M.A. Johnson, M. Lim, M.A. Gleeson, and J.M. Cregg. 2001. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. *Gene* 263:159-169.
8. Payne, W.E., C.A. Kaiser, B.J. Bevis, J. Soderholm, D. Fu, I.B. Sears, and B.S. Glick. 2000. Isolation of *Pichia pastoris* genes involved in ER-to-Golgi transport. *Yeast* 16:979-993.
9. Rhoads, D.M., C.S. Levings, 3rd, and J.N. Siedow. 1995. URF13, a ligand-gated, pore-forming receptor for T-toxin in the inner membrane of cms-T mitochondria. *J. Bioenerg. Biomembr.* 27:437-445.
10. Rossanese, O.W., C.A. Reinke, B.J. Bevis, A.T. Hammond, I.B. Sears, J. O'Connor, and B.S. Glick. A role for actin, Cdc1p, and Myo2p in the inheritance of late Golgi elements in *Saccharomyces cerevisiae*. *J. Cell Biol.* 153:47-61.
11. Rossanese, O.W., J. Soderholm, B.J. Bevis, I.B. Sears, J. O'Connor, E.K. Williamson, and B.S. Glick. 1999. Golgi structure correlates with transitional endoplasmic reticulum organization in *Pichia pastoris* and *Saccharomyces cerevisiae*. *J. Cell Biol.* 145:69-81.
12. Rothstein, R. 1991. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol.* 194:281-301.
13. Sears, I.B., J. O'Connor, O.W. Rossanese, and B.S. Glick. 1998. A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. *Yeast* 14:783-790.
14. Séron, K., V. Tieaho, C. Prescianotto-Baschong, T. Aust, M.-O. Blondel, P. Guillaud, G. Devilliers, O.W. Rossanese et al. 1998. A yeast t-SNARE involved in endocytosis. *Mol. Biol. Cell* 9:2873-2889.
15. Waterham, H.R., M.E. Digan, P.J. Koutz, S.V. Lair, and J.M. Cregg. 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and characterization of its promoter. *Gene* 186:37-44.

Thanks to Nathalie Glab for providing plasmid pEMBLye30/2-1urf13-TW, to Frances Henze of DuPont Agricultural Products for providing Lannate LV, and to the members of the Glick lab for helpful discussion. This work was supported by NSF grant no. MCB-9604342, NIH grant no. GM61156, and an award from the Pew Charitable Trusts. Address correspondence to Dr. Benjamin S. Glick, Department of Molecular Genetics and Cell Biology, The University of Chicago, 920 East 58th Street, Chicago, IL 60637, USA. e-mail: bsglick@midway.uchicago.edu

Received 7 November 2000; accepted 21 March 2001.

Jon Soderholm, Brooke J. Bevis, and Benjamin S. Glick
The University of Chicago
Chicago, IL, USA