Construction of a gene trap vector, pPTR-EGFP1, for the filamentous fungus, Aspergillus oryzae

Research Note

Construction of a gene trap vector, pPTR-EGFP1, for the filamentous fungus, *Aspergillus oryzae*

Satoshi SUZUKI, Ken-ichi KUSUMOTO and Yutaka KASHIWAGI
National Food Research Institute

Abbreviation: *egfp*, enhanced green fluorescent protein gene; REMI, restriction enzyme mediated integration; *ptrA*, pyrithiamine resistant gene of *Aspergillus oryzae*; SA, splice acceptor sequence; sCtm, terminator sequence of *A. nidulans sC* gene; *Amp*^{r}, ampicillin resistance gene; RACE, Rapid Amplification of cDNA ends; *DUR3*, urea active transporter gene of *Saccharomyces cerevisiae*; PCR, polymerase chain reaction; PEG, Polyethylene Glycol; DIG, digoxigenin

Abstract

A gene trap vector for *Aspergillus oryzae* was constructed for use in comprehensive analysis of gene function. This vector, pPTR-EGFP1, has the *egfp* gene as a reporter gene preceded by a splice acceptor sequence, followed by *A. nidulans sC* gene terminator and pyrithiamine resistant gene (*ptrA*) as a selective marker in *A.oryzae*. This vector was integrated into a host genome randomly to create tag-line transformants. Approximately 300 pyrithiamine-resistant transformants were thus obtained, one of which trapped the 5' end of the coding region of yeast *DUR3* homologue. Our results showed that pPTR-EGFP1 could be a useful tool in the gene function analysis of *A.oryzae*, namely, by isolating unknown genes by plasmid rescue and by trapping the endogenous promoter activity.

(Received Dec. 2, 2002 ; Accepted Feb. 5, 2003)

Filamentous fungus *Aspergillus oryzae* is one of the most important microorganisms in the Japanese fermentation industry, where it is used in production of sake, miso, and soy sauce, for example. In 2002, the consortium of the genome project of *A. oryzae* announced that they had assembled draft sequences of the entire genome.

In any post genome-sequence study, functional analysis of unknown genes is essential. Because neither the sexual reproduction cycle nor the genetic markers of *A. oryzae* have been identified or established, insertion mutagenesis is an effective tool in the gene function analysis of *A. oryzae*. Gene trapping is an effective method for simultaneously creating and tagging mutated genes.[1] In gene trapping, vectors that contain reporter genes or a selection marker are integrated into the genome randomly, and the phenotype of the mutants should reflect the alteration of the gene by integration of the vector.

In promoter trap and gene trap methods, the promoterless reporter gene or selection marker is connected directly to the endogenous promoter or coding region by vector integration into the genomic DNA. Gene trapping has been successfully applied to insertion mutagenesis of mice ES cells.[2, 3]

In this study, for comprehensive gene function analysis of *A. oryzae*, we constructed a gene trap vector for *A. oryzae* that can be applied to establish the gene disruption mutant library. A gene trap vector, pPTR-EGFP1 (accession number AB076373), was constructed (Fig.1). To confirm if the vector could tag unknown genes or trap endogenous promoter activity, the vector was then integrated into the

Corresponding author: Satoshi Suzuki, E-mail: satosuz@nfri.affrc.go.jp

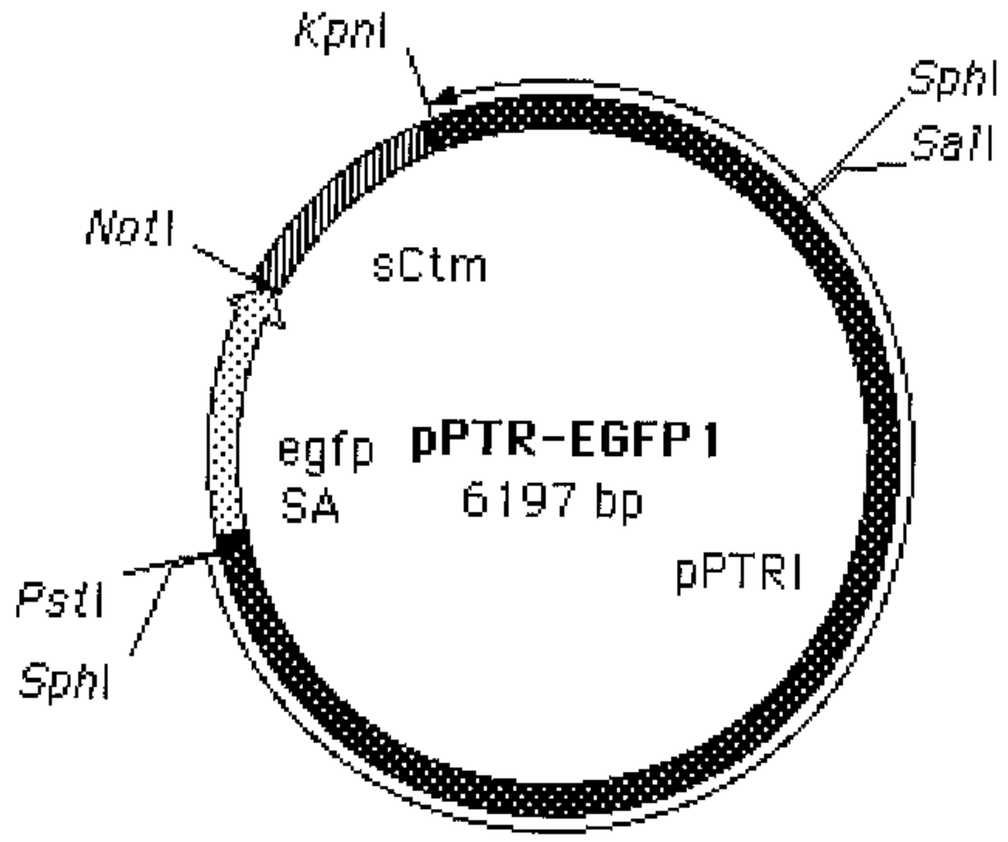

Fig.1 Construction of the gene trap vector, pPTR-EGFP1

SA and black box indicate splice acceptor sequence; egfp and dotted box indicate reporter gene egfp ; sCtm and striped box indicate terminator sequence of A. nidulans *sC* gene; pPTR I and checked box indicate pyrithiamine-resistant vector pPTR I (TAKARA).

genomic DNA of *A. oryzae* NFRI 1599. To confirm the integration of the vector pPTR-EGFP1, Southern blot analysis was done. To determine the expression of the reporter gene *egfp*, Northern blot analysis was done. Finally, to determine the mode of vector integration in the genome, the DNA fragment flanking the integrated vector was rescued by using the plasmid rescue method.

The vector was constructed as follows. This vector consisted of three parts: reporter gene *egfp* with a splice acceptor sequence, predicted terminator sequence of *A. nidulans sC* gene[4], and pyrithiamine-resistant vector pPTR I (TAKARA). Both the *egfp* gene with splice acceptor and the *sC* terminator sequence were amplified by PCR. The EGFP coding sequence that is preceded by the splice acceptor sequence was amplified by two rounds of PCR. The first round involved three components: pEGFP-N1 (Clontech) as a template, the sense primer EGF5'-2 (5'-ATGGTGAGCAAGGGCGAGGA-3'), and the antisense primer EGF3'-2 (5'-GCGGCCGCTTTACTTGTACA-3'). The second round involved using the first-PCR product as a template, the sense primer EGF5'-4 (5'-CTAACGCTATTCCAGATGGTGAGCA-3', where the underlined nucleotides indicate the splice acceptor sequence), and the antisense primer EGF3'-2. In the second round, the predicted splice acceptor sequence of *A. nidulans* was added to the *egfp* coding region. The resulting second-PCR product was then ligated to the pCR2.1 TOPO vector (Invitrogen). The gene of EGFP with splice acceptor (SA-EGFP) was then connected to the predicted terminator sequence of *A. nidulans sC* gene. This predicted *sC* terminator (sCtm) was then amplified by PCR, which involved pUSC[5,6] as a template and the sense primer sCT5'(5'-ATAAGAATGCGGCCGCATAATATTGGGCGTTAGAGATTTGCGATTT-3'), and the antisense primer sCT3'(5'-CCCAAGCTTAAGCTTCTCTTGGCAATAGCTGCCCGTATG-3'). The underlined nucleotides in this sense primer indicate non-homologous sequences added to create a restriction site *Not* I at the end of the PCR fragments. Then, the resulting PCR product was ligated to pCR2.1 TOPO (Invitrogen). The sCtm fragment was digested by the *Kpn* I and *Pst* I (restriction enzymes) sites of pCR2.1 and then ligated into the *Kpn* I and *Pst* I sites of pGEM-3zf, yielding pGEM-3zf-sCtm. SA-EGFP fragment was digested by the *Not* I site of the insert and by the *Pst* I site of pCR2.1 TOPO vector and ligated into pGEM-3zf-sCtm by the *Not* I and *Pst* I sites of pGEM-3zf-sCtm. The result as the plasmid pGEM-SA-EGFP-sCtm.

This plasmid was then digested with *Pst* I and *Kpn* I. The resulting SA-EGFP-sCtm fragment was ligated the *Pst*

I and *Kpn* I sites of pPTR I to yield pPTR-EGFP1. (Fig.1) The EGFP over-expression vector for positive-control strain was constructed by ligation of *TEF1* promoter[7] into the *Pst* I site of pPTR-EGFP1. The *TEF1* promoter was isolated from genomic DNA of *A.oryzae* NFRI 1599 (RIB40) by PCR.

To confirm if pPTR-EGFP1 could trap either the endogenous promoter activity or a 5' end of coding region, we integrated pPTR-EGFP1 into genomic DNA of *A. oryzae* NFRI 1599 by REMI using *Pst* I . Transformations of *A. oryzae* were done as described by Iimura et al.,[8] with slight modification as follows. In REMI, before the transformation, 20 μ g of the plasmid DNA was digested with the restriction nuclease *Pst* I . Strain NFRI 1599 conidia were cultured in 100ml of Czapek-Dox medium (CD medium) for 48hr at 30℃. The resulting mycelium were incubated with YATALASE (TAKARA) to obtain protoplasts according to the manufacturer's (TAKARA) instructions. Then, the protoplasts were incubated with the linear plasmid DNA and 100units of *Pst* I , and PEG solution was added at a final concentration of 33%. Transformants were selected on a CD medium agar plate containing 0.1 μ g/ml of pyrithiamine. The integration yielded 300 pyrithiamine-resistant transformants. Among the resulting 300 pyrithiamine-resistant transformants, 19 were putative fluorescent transformants (as revealed by fluorescence microscopy). For these 19 transformants, a cover-slip culture method was used to discriminate between EGFP fluorescence and autofluorescence of dead cells. The results suggested that only one strain (strain 23) was a putative EGFP-expressing transformant.

To confirm the integration of the vector pPTR-EGFP1, Southern blot analysis was done for 5 of the 19 putative fluorescent transformants (strains 23, 31, 62, 64, and 85) selected at random. General techniques for nucleic acid manipulation were done according to the standard methods described by Sambrook et al.[9] Southern analysis was done on *Hpa* I and *Kpn* I digested genomic DNA samples from the five strains using the vector sequence as a probe. The restriction enzyme *Hpa* I has no recognition site and *Kpn* I has one recognition site in the vector. In all five strains, the *Hpa* I -digested genomic DNA samples displayed a single hybridization band of different sizes, suggesting that the vector integration occurred at random positions via illegitimate recombination at one site in the genome. In strains 23 and 31, the *Kpn* I -digested genomic DNA samples displayed several hybridization bands, suggesting that the vector integration occurred by tandem

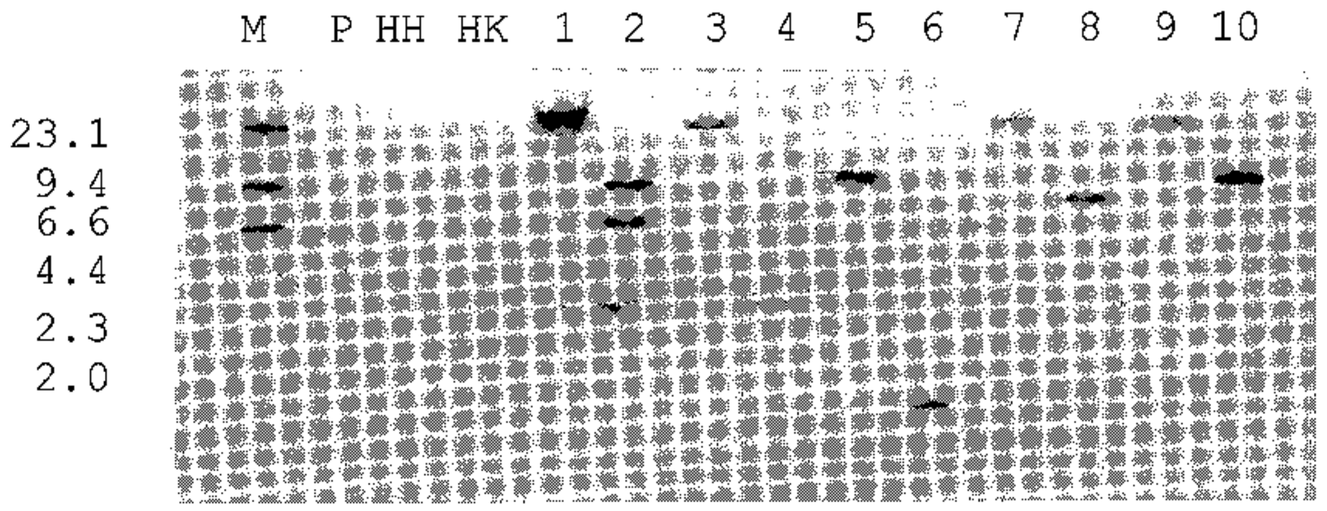

Fig.2 Southern blot analysis of putative fluorescent transformants

M, Dig-labeled DNA size marker (λ.../Hind Ⅲ)(Roche); P, positive-control (1ng of linearized pPTR-EGFP1); HH, *Hpa* I -digested genomic DNA of the host strain NFRI 1599; HK, KpnáT-digested genomic DNA of strain NFRI 1599; Lanes 1 and 2, strain 23; lanes 3 and 4, strain 31; lanes 5 and 6, strain 62; lanes 7 and 8, strain 64; lanes 9 and 10, strain 85. The odd- or even-numbered lanes represent the *Hpa* I or *Kpn* I digested genomic DNAs of the transformants, respectively. Each lane contained 2 μ g of genomic DNA.

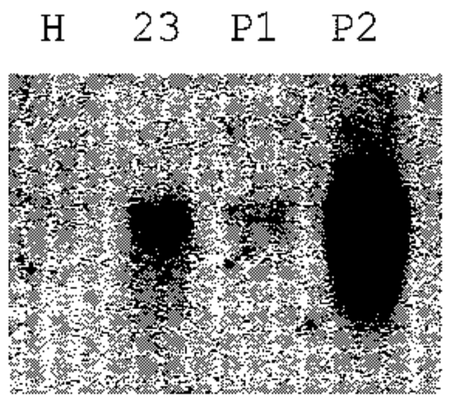

Fig.3 Northern blot analysis of strain 23, a putative EGFP-expressing transformant

H, 5μg mRNA of strain NFRI 1599; 23, 5μg mRNA of strain 23; P1, 0.5μg mRNA of *egfp* over-expressed strain; P2, 5μg mRNA of *egfp* over-expressed strain.

integration at a single site (Fig.2). In strains 62, 64, and 85, the *Kpn* I -digested genomic DNA samples displayed a single hybridization band, suggesting one-copy integration (Fig.2).

To determine the expression of the reporter gene *egfp*, Northern blot analysis was done for strain 23 (Fig.3).

The transcript size of the *egfp* gene in strain 23 was similar to that in the positive-control strain, namely, over-expression of the *egfp* by *TEF1* promoter constitutively. This suggests that this transcript in strain 23 was not a fusion transcript of *egfp* with an endogenous gene. If vector integration had occurred in the intron sequence of an endogenous gene, then the transcript size would have increased. An endogenous exon and *egfp* gene are fused in an in-frame manner during the RNA-splicing via the splice acceptor sequence in the vector. The *egfp* expression level of strain 23 was found comparable to that of the control strain.

Finally, to determine the mode of vector integration in the genome, the 9kb genomic DNA fragment flanking the integrated vector from strain 23 was recovered by using the plasmid rescue method. In this rescue, 4μg of strain 23 genomic DNA were digested by *Hpa* I , and then the resulting genomic fragments were self-ligated by using a DNA Ligation Kit Ver.1 (TAKARA). The ligated and circularized DNA were transformed into *E. coli* DH10B by electroporation, and the resulting transformants of *E. coli* were spread on a LB/ampicillin selection agar plate. Each

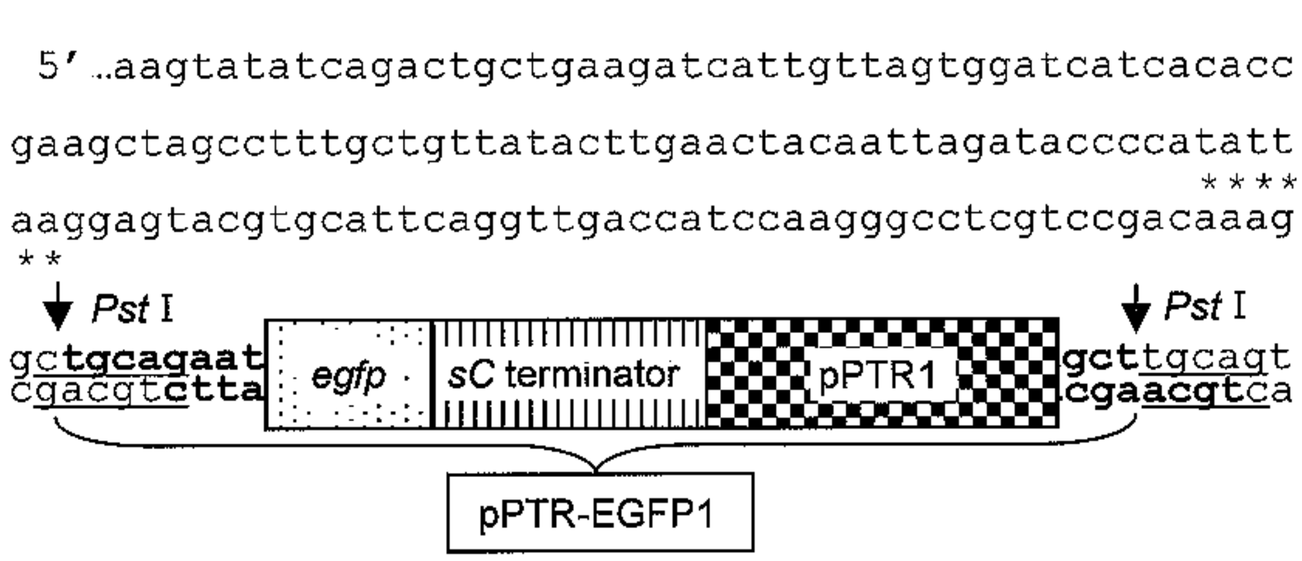

Fig.4 Nucleotide and deduced amino acid sequences of the flanking regions of the vector integration site in the genomic DNA of strain 23

The 5' and 3' flanking genome sequences adjacent to the plasmid integration site are indicated by non-bold letters and the plasmid sequence is indicated by bold letters. The upper-case letters indicate the coding region and the lower-case letters indicate non-coding region and putative intron sequences. ↓ indicates the cut-off point of genomic DNA by *Pst* I during REMI. Underlined "ctgcag" indicates recognition sequence of *Pst* I . * indicates putative TATA sequences. The square box indicates the structure of the vector (pPTR-EGFP1). The 3' flanking genome region encodes the homologue of Saccharomyces cerevisiae *DUR3*, and deduced amino acid sequence is indicated below the nucleotide sequence of the 5' end of the coding region of *DUR3* homologue.

1kb sequences of the 5' and 3' flanking regions of the integrated plasmid vector was determined, and the results are shown in Fig. 4. Homology search of the 3' flanking sequence using BLAST on the NCBI website (http://www.ncbi.nlm.nih.gov/) revealed that plasmid integration occurred near the 5' end of a gene homologous to yeast *DUR3*[10], which is the urea active transporter gene. The deduced amino acid sequences from nucleotide sequences of the homologue of *DUR3* are indicated in Fig.4. One hypothesis for the plasmid integration is that the genomic DNA of strain 23 is cut at the *Pst* I site existing between the 5' regulatory sequence and coding region of the homologue of *DUR3*, and then pPTR-EGFP1 is integrated into this *Pst* I site during REMI event (Fig.4)

In this study, we constructed a gene trap vector, pPTR-EGFP1, and showed its usefulness not only for tagging unknown genes but also for trapping endogenous promoter activity. This achievement is the first step in constructing the gene disruption mutant library of *A. oryzae*. Although this gene trap vector might be less effective for finding promoter activity than the "promoter probe" strategy described by Luo[11] and Ozeki et al.[12], it allows direct isolation of unknown genes by using plasmid rescue or 5' RACE.

Based on our results, the gene trap vector, pPTR-EGFP1, offers an effective tool for isolating an unknown gene by plasmid rescue and by trapping the endogenous promoter activity.

Acknowledgements

This work was supported in part by a Grant-in-Aid (Bio Design Project) from the Ministry of Agriculture, Forestry and Fisheries Japan. We thank Prof. Dr. K. Gomi for useful discussions, and thank Dr. K. Iwashita for providing the pUSC.

References

1) mbrowicz, B.P. and Friedrich, G.A., Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse, *Int.J.Dev.Biol.*, **42**, 1025-1036 (1998).
2) Hicks, G.G., Shi, E.G., Li, X.M., Li, C.H., Pawlak, M. and Ruley, H.E., Functional genomics in mice by tagged sequence mutagenesis, *Nature Genet.*, **16**, 338-344 (1997).
3) Gossler, A., Joyner, A., Rossant, J., and Skarnes, W.C., Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes, *Science*, **244**, 463-465 (1989).
4) Buxton, F.P., Gwynne, D.I., Davies, R.W., Cloning of a new bidirectionally selectable marker for *Aspergillus* strains, *Gene*, 84,329-34 (1989).
5) Iwashita, K., Nagahara, T., Kimura, H., Takano, M., Shimoi, H., Ito, K., The bglA gene of *Aspergillus kawachii* encodes both extracellular and cell wall-bound beta-glucosidases, *Appl Environ Microbiol.*, **65**, 5546-5553 (1999).
6) Yamada, O., Lee, B.R., Gomi, K., Transformation System for *Aspergillus oryzae* with Double Auxotrophic Mutations, *niaD* and *sC. Biosci, Biotech. Biochem.*, **61**, 1367-1369 (1997).
7) Kitamoto, N., Matsui, J., Kawai, Y., Kato, A., Yoshino, S., Ohmiya, K., Tsukagoshi, N., Utilization of the TEF1- α gene (*TEF1*) promoter for expression of polygalacturonase genes, *pgaA* and *pgaB*, in *Aspergillus oryzae, Appl Microbiol Biotechnol.*, **50**, 85-92 (1998).
8) Iimura, Y., Gomi, K., Uzu, H., Hara, S., Transformation of *Aspergillus oyrzae* through Plasmid-mediated Complementation of the Methionine-auxotrophic Mutation, *Agric. Biol. Chem.* ,**51**, 323-328 (1987).
9) Sambrook,J., Fritsch,E.F., Maniatis,T., Molecular cloning: a laboratory manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).
10) Sumrada,R., Gorski,M., Cooper, T., Urea Transport-Defective Strains of *Saccharomyces cerevisiae, J Bacteriol.*, **125**, 1048-1056 (1976).
11) Luo, X., Cloning and characterization of three *Aspergillus niger* promoters, *Gene*, **163**, 127-131 (1995).
12) Ozeki, K., Kanda, A., Hamachi, M., Nunokawa, Y., Construction of Promoter Probe Vector Autonomously Maintained in Aspergillus and Characterization of Promoter Regions Derived from *A. oryzae* Genomes, *Biosci. Biotech. Biochem.*, **60**, 383-389 (1996).

麹菌用ジーントラップベクター，pPTR-EGFP1の作成

鈴木　聡・楠本 憲一・柏木　豊

麹菌 *Aspergillus oryzae* の包括的な遺伝子機能解析に資するため、新規のジーントラップ遺伝子破壊用プラスミドベクターを開発した。このベクター、pPTR-EGFP1 はスプライス受容保存配列と *A. nidulans* の *sC* 遺伝子ターミネーター配列を付加した緑色蛍光タンパク質遺伝子をレポーター遺伝子として保持し、麹菌で働く選抜マーカー遺伝子としてピリチアミン耐性遺伝子（*ptrA*）を保持する。このベクターを宿主麹菌のゲノム DNA 内にランダムに挿入することにより、タグライン形質転換株を作成することができる。このベクターを用いた形質転換実験により、約 300 株のピリチアミン耐性株が得られ、そのうちの 1 株が酵母の尿素運搬体タンパク遺伝子に相同な遺伝子のコード領域の 5’末をトラップしていた。本ベクター pPTR-EGFP1 を用いることにより、形質転換株からのプラスミドレスキュー法による未知遺伝子の回収、及び内在性のプロモーター活性のトラップが可能となることが示され、麹菌 *A. oryzae* の遺伝子機能解析において pPTR-EGFP1 が有効に機能することが明らかとなった。