Current Genetics (2019) 65:785–798
https://doi.org/10.1007/s00294-019-00934-6

ORIGINAL ARTICLE

Basal transcription profiles of the rhamnose-inducible promoter P_{LRA3} and the development of efficient P_{LRA3}-based systems for markerless gene deletion and a mutant library in *Pichia pastoris*

Jian Jiao[1] · Shuai Wang[1] · Mingli Liang[1] · Yuhong Zhang[1] · Xinxin Xu[1] · Wei Zhang[1] · Bo Liu[1]

Received: 21 November 2018 / Revised: 17 December 2018 / Accepted: 8 January 2019 / Published online: 24 January 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

An ideal inducible promoter presents inducibility with an inducer and no basal transcription without inducer. Previous studies have shown that P_{LRA3} in *Pichia pastoris* is a strong rhamnose-inducible promoter for driving the industrial production of recombinant proteins. However, another important profile of P_{LRA3}, the basal transcription, was not investigated yet. In this study, the basal transcription of P_{LRA3} was assessed according to the profiles of two test strains grown in media lacking rhamnose: (1) the production of secretory β-galactosidase in *P. pastoris* GS115/P_{LRA3}-LacB, in which *lacB* expression was regulated by $P_{LRA3,}$ and (2) growth in *P. pastoris* GS115/P_{LRA3}-MazF, in which the expression of *mazF*, which encodes an intracellular toxic protein, was controlled by P_{LRA3}. Analyses revealed low β-galactosidase production and non-obviously inhibited growth of the test strains, which suggests that there was a low basal transcription level of P_{LRA3} when rhamnose was absent. Thus, P_{LRA3} was an excellent candidate for genetic manipulation in *P. pastoris* due to its strict regulation, a strong and a low transcriptional activity with and without rhamnose, respectively. Subsequently, two systems were developed based on P_{LRA3} in *P. pastoris*: (1) an efficient markerless gene deletion system for single or multiple genes and (2) a high efficient *piggyBac* transposase-mediated mutation system for investigating the functions of unknown genes, as well as for the screening of expected mutants.

Keywords *Pichia pastoris* · Rhamnose-inducible promoter · Basal transcription · Markerless gene deletion · Mutation system

Introduction

The precise control of the expression levels of desired genes is essential for basic biological research and biotechnological applications. Inducible promoters allow for the transcription of target genes to be switched on or off in the presence or absence of inducers (Meisner and Goldberg 2016). Therefore, inducible promoters are widely used for theoretical studies, as well as for the industrial production of target proteins. A variety of inducible promoters, including the methanol-inducible *AOX1* promoter from *Pichia pastoris* (Unver et al. 2018; Zhao et al. 2018), rhamnose-inducible *rhaBAD* promoter from *E. coli* (Meisner and Goldberg 2016), galactose-inducible *GAL* promoter from *S. cerevisiae* (Ahn et al. 2013), and nitrate-inducible *nitA* promoter from *R. rhodochrous* (Herai et al. 2004), have been used for the industrial production of desired proteins. Additionally, a variety of inducible promoters can be applied to facilitate different types of molecular genetic manipulation: for example, the control of specific gene expression levels in metabolic engineering for improving the production of target products (Da Silva and Srikrishnan 2012; Trassaert et al. 2017), the development of markerless deletion systems (Yu et al. 2008; Luo et al. 2016), investigations of the

Communicated by M. Kupiec.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00294-019-00934-6) contains supplementary material, which is available to authorized users.

✉ Wei Zhang
zhangwei02@caas.cn

✉ Bo Liu
liubo01@caas.cn

[1] Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, No. 12 Zhongguancun South Street, Beijing 100081, China

connections between health and disease associated with gut microbiota (Lim et al. 2017), and the improved tetracycline-regulatable systems for conditional expression of genes in *C. albicans* and *C. tropicalis* (Bijlani et al. 2018).

Typically, an ideal promoter should be completely repressed in the absence of an inducer, as well as tunable to different transcription levels using various concentrations of an inducer (Rodriguez-Garcia et al. 2005). However, many inducible promoters exhibit obvious levels of background expression due to leaky expression and, as a result, target proteins, especially growth-inhibited or toxic proteins, are synthesized under non-inducible conditions and can result in host growth defects or even cell death (Giacalone et al. 2006). Thus, it is important to use tightly regulated promoters with low levels of leaky expression that enable host strains to grow normally until induction to produce target proteins, particularly when toxic proteins are produced (Saida et al. 2006).

A previous study from our research group showed that P_{LRA3} is a strong promoter that efficiently drives the production of desired proteins in *P. pastoris* using rhamnose as the inducer (Liu et al. 2016; Yan et al. 2018). However, due to leaky expression, the precise nature of the basal transcription of P_{LRA3} has yet to be fully elucidated. Thus, the present study evaluated the basal transcription profiles of P_{LRA3} in the absence of rhamnose induction. The results indicated that P_{LRA3} exhibited low basal transcription levels and has the potential for molecular genetic manipulation in *P. pastoris*. Subsequently, based on P_{LRA3}, a markerless gene deletion system and a mutation system were developed in *P. pastoris*. The two systems were highly efficient and may be excellent tools for molecular genetic manipulation in *P. pastoris*.

Materials and methods

Strains and media

The present study used *E. coli* Trans1-T1 (TransGen; Beijing, China) as the gene cloning host, *P. pastoris* GS115 (Invitrogen; Carlsbad, CA, USA) as the expression host, and the *pEasy*-Blunt Simple Cloning Vector (TransGen). The plasmids (pGH01, pGHLRA3, and pGHGAP/*gfp*) and *lacB*-expressing strains (*P. pastoris* GS115/P_{LRA3}-LacB, *P. pastoris* GS115/P_{AOX1}-LacB, and *P. pastoris* GS115/P_{GAP}-LacB) used in the present study have been described previously (Liu et al. 2016).

The MD, MR, and MM media contained 300 mM potassium phosphate, 1.34% YNB, 0.00004% biotin, and either 2% dextrose for MD, 2% rhamnose for MR or 1% (v/v) methanol, which was added every day, for MM.

The MDH, MRH, and MMH media contained 300 mM potassium phosphate, 1.34% YNB, 0.00004% biotin, 0.004% histidine, and either 2% dextrose for MDH, 2% rhamnose for MRH, or 1% (v/v) methanol, which were added every day, for MMH.

The YPR, YPD and YPM media contained 1% yeast extract and 2% peptone, and either 2% rhamnose for YPR, 2% dextrose for YPD, or 1% (v/v) methanol, which were added every day, for YPM.

The MDH medium was supplemented with leucine to a final concentration of 0.004% for preparation of the MDHL medium. The MDH, MRH and YPD media were supplemented with Zeocin to a final concentration of 100 μg/ml for preparation of the MDHZ, MRHZ and YPDZ media, respectively. To prepare the solid media, agar was supplemented into the above-mentioned media to a final concentration of 2% (w/v).

The primers used in this study are listed in Supplementary Table S1.

Determination of *lacB* expression and β-galactosidase activity

The single colonies of different recombinant *Pichia* strains (*P. pastoris* GS115/P_{LRA3}-LacB, *P. pastoris* GS115/P_{AOX1}-LacB, and *P. pastoris* GS115/P_{GAP}-LacB with *P. pastoris* GS115 as a control strain) was inoculated into the YPD medium and cultured at 28 °C for 36 h while shaking (200 rpm). Next, the cultures were inoculated at 0.1% (v/v) into fresh MDH, MMH, YPD, and YPM media, respectively, and then cultured at 28 °C while shaking. The cultures were sampled at 12-h intervals to determine β-galactosidase activity in the fermentation supernatants using the following procedure.

First, 800 μl of 0.25% (w/v) ortho-nitrophenyl-β-galactoside (*o*NPG) in phosphate–citrate buffer (50 mM, pH 5.2) was pre-incubated at 60 °C for 5 min. Then, 200 μl of enzyme solution was added and incubated at 60 °C for 15 min. Next, 1 ml of 10% trichloroacetic acid and 2 ml of 1 M Na_2CO_3 were sequentially added into the mixture to stop this reaction for the chromogenic reaction, and absorbance was measured at 420 nm. One unit of β-galactosidase was defined as the amount of enzyme that released 1 μmol of *o*-nitrophenol per minute under standard conditions (pH 5.2, 60 °C, 15 min).

Construction and growth profiles of *P. pastoris* harboring P_{LRA3}-MazF

Genomic DNA from *P. pastoris* was isolated using the TIANamp Yeast DNA Kit (Tiangen), and the left and right homologous DNA fragments of the sorbitol dehydrogenase gene (*sdh*) were cloned from the genomic DNA by two pairs of primers, *sdh*-F1/*sdh*-R1 and *sdh*-F2/*sdh*-R2, respectively. The two DNA fragments were fused into a single DNA

fragment named "SDH" via overlap-extension polymerase chain reaction (PCR) with the primers *sdh*-F1 and *sdh*-R2. The SDH DNA fragment was ligated into the *pEasy*-Blunt Simple Cloning Vector to generate plasmid pPICS01, which could be digested by *Sna*B I due to the introduction of an additional restriction enzyme site for the insertion of a DNA fragment between the left and right homologous recombinant DNA fragments.

The toxic endoribonuclease MazF belonged to the members of Type II toxin–antitoxin systems from *E. coli* (Berghoff and Wagner 2017), and was used to assess the basal transcription level of P_{LRA3}. The nucleotide sequence of the coding region of *mazF* (https://www.ncbi.nlm.nih.gov/gene/947252) was cloned from *E. coli* JM109. Next, *mazF* was ligated into plasmid pGHLRA3 via two restriction enzyme sites, *Eco*R I and *Not* I, which generated the *mazF* expression cassette in which *mazF* expression was controlled by P_{LRA3} as the promoter and *AOX1TT* as the terminator. Simultaneously, the *zeocin* expression cassette was amplified from plasmid pPICZA by PCR with a pair of primers, Zeocin-F and Zeocin-R. Subsequently, a tandem expression cassette (*mazF*–*zeocin*) containing the *mazF* expression cassette and the *zeocin* expression cassette was constructed via overlap-extension PCR and inserted into plasmid pPICS01 at the restriction enzyme site of *Sna*B I to generate plasmid pPICS*mazF*–zeo. After linearization with *Swa* I, the plasmid was introduced into *P. pastoris* GS115 via electroporation and the tandem expression cassette *mazF*–*zeocin* was integrated into chromosomal DNA at the *sdh* locus. The positive transformants, which were designated as *P. pastoris* GS115/P_{LRA3}-MazF, were screened onto YPD with Zeocin (100 µg/ml) and then identified using PCR.

Pichia pastoris GS115/P_{LRA3}-MazF, along with *P. pastoris* GS115 and *P. pastoris* GS115/P_{LRA3}-LacB as control strains, was independently cultured in YPD at 28 °C for 48 h while shaking (200 rpm). Optical density at 600 nm (OD_{600}) was measured in the cultures and adjusted to the same value using sterile water. Next, the cultures were inoculated at 0.1% (v/v) using fresh MDH and MRH media, respectively, and cultured at 28 °C while shaking. The cultures were sampled at 12-h intervals to determine OD_{600}.

Markerless gene deletion plasmid pHISABCZ–*mazF*

For the markerless deletion of *his4*, a 1.2-kb left homologous sequence (A) and a 0.6-kb right homologous sequence (C) were amplified from chromosomal DNA via PCR with the following pairs of primers: *his4A*-F/*his4A*-R and *his4C*-F/*his4C*-R, respectively; there was a 29-bp overlap between *his4A*-R and *his4C*-F. Then, two restriction enzyme sites, *Pme* I and *Sna*B I, were introduced and DNA fragments A and C were fused into a single DNA fragment designated as HIS4AC, which was ligated into the *pEasy*-Blunt Simple Cloning Vector to generate plasmid pHISAC01. Additionally, a 0.6-kb DNA fragment (B) upstream of the 3′ end of DNA fragment A was amplified from chromosomal DNA via PCR and inserted into plasmid pHISAC01 at restriction site *Sna*B I to develop plasmid pHISABC01. Finally, the *mazF*–*zeocin* expression cassette was inserted into pHISABC01 at restriction site *Pme* I to generate the markerless deletion plasmid pHISABCZ–*mazF*.

Construction and identification of markerless gene deletion strains

After linearization with *Swa* I, plasmid pHISABCZ–*mazF* was introduced into *P. pastoris* GS115 and integrated into chromosomal DNA at the *his4* locus. The positive transformants, designated as *P. pastoris* GS115/*his4*Z–*mazF*, were screened onto YPDZ and then identified using PCR.

Pichia pastoris GS115/*his4*Z–*mazF* was inoculated into YPD and incubated at 28 °C for 36 h while shaking. The cultures were then inoculated into the YPD and YPR media at 0.1% (v/v) each and grown at 28 °C while shaking. The OD_{600} values of the cultures were measured at 12-h intervals. Simultaneously, the cultures were each diluted with sterile water and plated onto solid YPD medium until single colonies occurred. Approximately 100 single colonies were transferred onto each of the YPD and YPDZ media, and the number of big colonies on the YPD medium and number of small colonies on the YPDZ medium were counted. PCR was performed using chromosomal DNA from the cells of small colonies as templates, and then the PCR products were sequenced to assay whether the target gene was deleted.

Construction of plasmid p53TRZ/P_{LRA3}-PB

A DNA fragment containing the inverted terminal repeats of *piggyBac* (5TR/3TR) recognized by *piggyBac* transposase (GenBank: EF587698.1) and multiple cloning sites (MCS) was synthesized (Table S2). The DNA fragment was inserted into the *pEasy*-Blunt Simple Cloning Vector to generate the plasmid p53TR. Then, the *zeocin* expression cassette was cloned from plasmid pPICZA and inserted into plasmid p53TR at restriction enzyme site *Pme* I to develop plasmid p53TRZ.

The sequence of *piggyBac* encoding PB was synthesized by GenScript Biotech Corp in Nanjing, China. Then, *piggyBac* was inserted into plasmid pGHLRA3 via *Eco*R I and *Not* I to generate the resultant plasmid pGHLRA3/P_{LRA3}-PB, in which *piggyBac* expression was under the control of P_{LRA3}. Plasmid pGHLRA3/P_{LRA3}-PB was digested by *Asc* I and *Pac* I and the DNA fragment containing *piggyBac* was ligated into plasmid p53TRZ after the digestion of *Asc* I and *Pac* I to generate the plasmid p53TRZ/P_{LRA3}-*PB*.

Development of the leucine-deficient strain *P. pastoris* GS115/*leu2T*

The expression cassette of *leu2* was amplified from the genomic DNA of *P. pastoris* cells by PCR with the following two pairs of primers: *leu2*-F1/*leu2*-R1and *leu2*-F2/*leu2*-R2. Using the above two PCR products and plasmid p53TRZ/P_{LRA3}-*PB* as the templates and *leu2*-F1 and *leu2*-R2 as the primers for overlap-extension PCR, the PCR products were introduced into *P. pastoris* GS115 and then integrated into the *leu2* locus via homologous recombination to develop the strain *P. pastoris* GS115/*leu2T*. This resultant strain, *P. pastoris* GS115/*leu2T*, was deficient in *leu2* due to the insertion of an exogenous DNA fragment into the coding sequence (CDS) of *leu2* and could not survive in leucine-deficient medium.

Mutation library of *P. pastoris* GS115

Pichia pastoris GS115/*leu2T* was grown into YPR or MRH media and produced PB following rhamnose induction at P_{LRA3}. Once PB was expressed, the exogenous DNA fragment would move from the *leu2* locus with four nucleotides (TTAA) remaining. These four nucleotides and two other nucleotides (GC) introduced into PCR primers in turn introduced two additional codons into the CDS of *leu2*. The strain designated as *P. pastoris* GS115/*leu2m* would then present with leucine-autotrophic growth. Subsequently, exogenous DNA would lose or insert into other TTAA sites, thereby mutating *P. pastoris* GS115; the mutants could be screened onto MDHZ or MRHZ media.

gfp expression and assays of green fluorescence intensity

Green fluorescent protein (GFP) gene *gfp* was obtained from plasmid pGHGAP/*gfp* via restriction enzymes *Eco*R I and *Not* I, and then was inserted into plasmid pPIC3.5 digested by *Eco*R I and *Not* I to develop plasmid pPIC3.5/*gfp*. Plasmid pPIC3.5/*gfp* after linearization with *Bgl* II was, respectively, introduced into *P. pastoris* GS115 and *P. pastoris* GS115/*leu2T* to develop strains *P. pastoris* GS115/P_{AOX1}-GFP and *P. pastoris* GS115/ *leu2T*/P_{AOX1}-GFP in which expression of *gfp* was regulated by P_{AOX1}. Strain *P. pastoris* GS115/P_{GAP}-GFP in which *gfp* expression was controlled by the strong constitute promoter P_{GAP} was used a positive control.

The above three strains, *P. pastoris* GS115/P_{AOX1}-GFP, *P. pastoris* GS115/*leu2T*/P_{AOX1}-GFP, and *P. pastoris* GS115/P_{LRA3}-GFP, were inoculated into the YPD medium and cultured at 28 °C for 36 h while shaking (200 rpm). Next, the cultures were inoculated at 0.1% (v/v) into fresh MRHZ media, respectively, and then cultured at 28 °C while shaking to OD_{600} reaching to ~ 2.0. Green fluorescence intensities from ~ 1,000,000 cells from each culture were monitored by cell analyzer (BD LSRFortessa).

Results

LacB expression profiles of *P. pastoris* GS115/P_{LRA3}-LacB under non-inducible conditions

Although P_{LRA3} can be activated to efficiently drive the expression of target genes via induction with rhamnose (Liu et al. 2016; Yang et al. 2018), its basal transcription level has yet to be fully elucidated. Thus, to investigate the basal transcription level of P_{LRA3}, the *P. pastoris* GS115/P_{LRA3}-LacB strain, in which *lacB* expression was controlled by P_{LRA3}, was developed. β-Galactosidase encoded by *lacB* can be secreted into the fermentation supernatants of cultures and, therefore, β-galactosidase activities can be used as an index of the transcription levels of P_{LRA3}. Thus, the basal transcription level of P_{LRA3} was preliminarily assayed according to β-galactosidase activity in the fermentation supernatants of *P. pastoris* GS115/P_{LRA3}-LacB grown under non-inducible conditions. Simultaneously, the *P. pastoris* GS115/P_{AOX1}-LacB and *P. pastoris* GS115/P_{GAP}-LacB strains, in which *lacB* expression levels were controlled by P_{AOX1} and P_{GAP}, respectively, were used as the controls.

The three strains, *P. pastoris* GS115/P_{LRA3}-LacB, *P. pastoris* GS115/P_{AOX1}-LacB, and *P. pastoris* GS115/P_{GAP}-LacB, were grown in non-inducible glucose-containing YPD and MD media and β-galactosidase activity in the fermentation supernatants was determined at various intervals (Fig. 1a, b). Very little β-galactosidase activity was detected in the fermentation supernatants of *P. pastoris* GS115/P_{LRA3}-LacB and *P. pastoris* GS115/P_{AOX1}-LacB. When estimated according to the specific activity of β-galactosidase (575 U/mg), the maximal amounts of β-galactosidase produced for *P. pastoris* GS115/P_{LRA3}-LacB and *P. pastoris* GS115/P_{AOX1}-LacB were 0.1 µg/ml and 0.017 µg/ml, respectively, in the YPD medium and 0.17 µg/ml and 0.012 µg/ml, respectively, in the MD medium. In the YPR and MR media containing inducible rhamnose, β-galactosidase production in the fermentation supernatants of *P. pastoris* GS115/P_{LRA3}-LacB reached approximately 34.8 µg/ml and 17.4 µg/ml, respectively. Thus, P_{LAR3} exhibited low basal transcription levels under non-inducible conditions.

Methanol is the most commonly used inducer for the production of heterologous proteins in *P. pastoris*; therefore, the basal transcription profiles of P_{LRA3} in methanol-containing YPM and MM media were also investigated. The maximal β-galactosidase production levels in the fermentation supernatants of *P. pastoris* GS115/P_{LRA3}-LacB cultured in YPM and MM media were approximately 0.17 µg/ml and 0.06 µg/

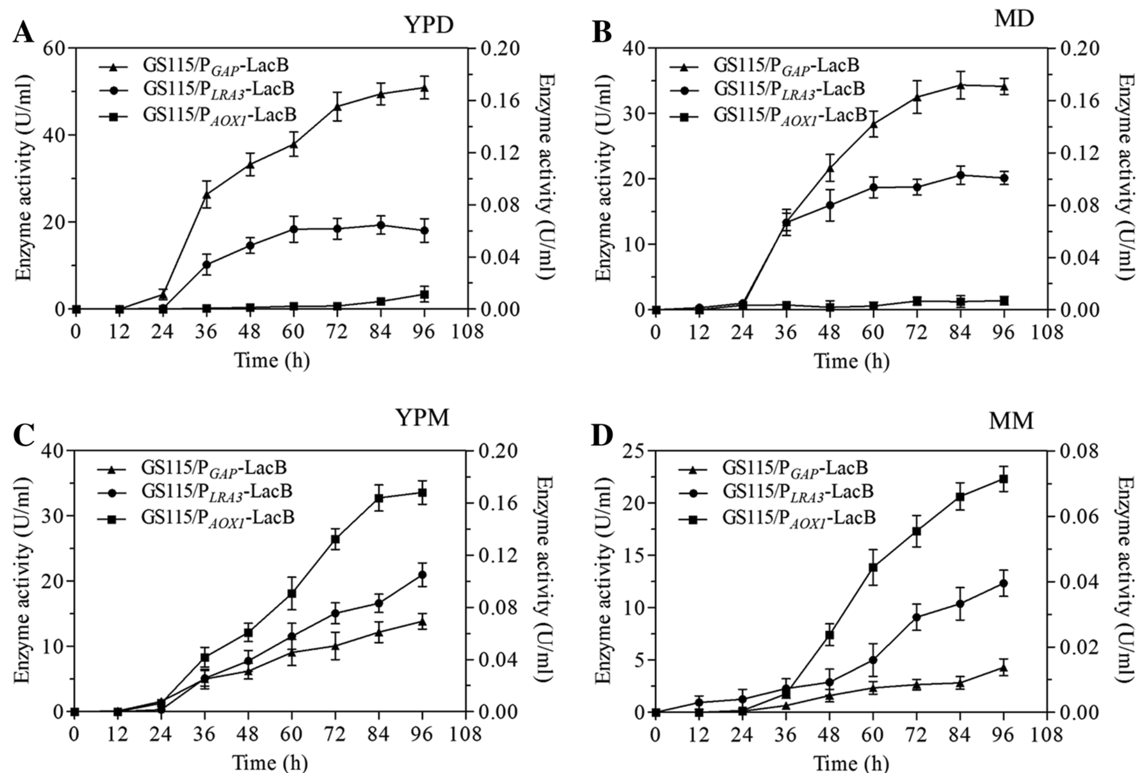

Fig. 1 Basal transcription assay of P_{LRA3} according to β-galactosidase activities in the fermentation supernatants in *P. pastoris* GS115/P_{LRA3}-LacB grown in YPD (**a**), MD (**b**), YPM (**c**), MM (**d**). Methanol-inducible promoter P_{AOX1} and constitute strong promoter P_{GAP} were used as the controls. *P. pastoris* GS115/P_{LRA3}-LacB, *P. pastoris* GS115/P_{AOX1}-LacB and *P. pastoris* GS115/P_{GAP}-LacB were the strains in which expression of *lacB* was under control of P_{LRA3}, P_{AOX1} and P_{GAP}, respectively. Each test was performed in triplicates, and the results are presented as means ± SD of three replicates

ml, respectively (Fig. 1c, d). The control strains *P. pastoris* GS115/P_{AOX1}-LacB and *P. pastoris* GS115/P_{GAP}-LacB produced massive amounts of β-galactosidase in the fermentation supernatants of the cultures (Fig. 1c, d). Taken together, these results suggest that P_{LRA3} had a low leaky transcription profile when methanol was used as the main carbon. Furthermore, P_{LRA3} was tightly regulated by rhamnose and exhibited relatively low basal transcription levels under non-inducible conditions even though it was less tightly regulated than the well-known methanol-inducible promoter P_{AOX1}.

Growth profiles of *P. pastoris* GS115/P_{LRA3}-MazF

A secretory protein, β-galactosidase, was used as the reporter to indicate the basal transcription of P_{LRA3}. Secretory proteins might not be fully secreted into the fermentation supernatants of cultures and non-secretory compartments are resident in cells; as a result, β-galactosidase activities in fermentation supernatants might not accurately reflect the basal transcription levels of P_{LRA3}. To account for this possibility, the non-secretory protein MazF was used as an additional reporter for assay of the basal transcription levels of P_{LRA3}. MazF is encoded by *mazF*, which is a sequence-specific endoribonuclease that can inhibit protein synthesis and lead to cell death. Theoretically, the growth profiles of *P. pastoris* GS115/P_{LRA3}-*mazF*, in which *mazF* expression is regulated by P_{LRA3} would be closely related to the production of MazF. Due to the massive production of MazF that arises from rhamnose induction at P_{LRA3}, *P. pastoris* GS115/P_{LRA3}-*mazF* could not survive in the YPR and MRH media containing rhamnose as a carbon source. When inoculated into YPD and MDH media containing glucose as the carbon source, *P. pastoris* GS115/P_{LRA3}-MazF grew normally if P_{LRA3} was tightly regulated by rhamnose, and died if P_{LRA3} exhibited some degree of basal transcription.

Several experiments were performed to determine if the above assumptions were correct. *P. pastoris* GS115/P_{LRA3}-MazF and two control strains, *P. pastoris* GS115 and *P. pastoris* GS115/P_{LRA3}-LacB, grew well on the YPD and MDH media (Fig. 2a, c), which indicated that the growth of *P. pastoris* GS115/P_{LRA3}-MazF was not inhibited by low or trace amounts of MazF due to the low basal transcription of P_{LRA3}. When inoculated into rhamnose-containing YPR and MRH media, *P. pastoris* GS115/P_{LRA3}-MazF did not survive due to the massive production of MazF following rhamnose

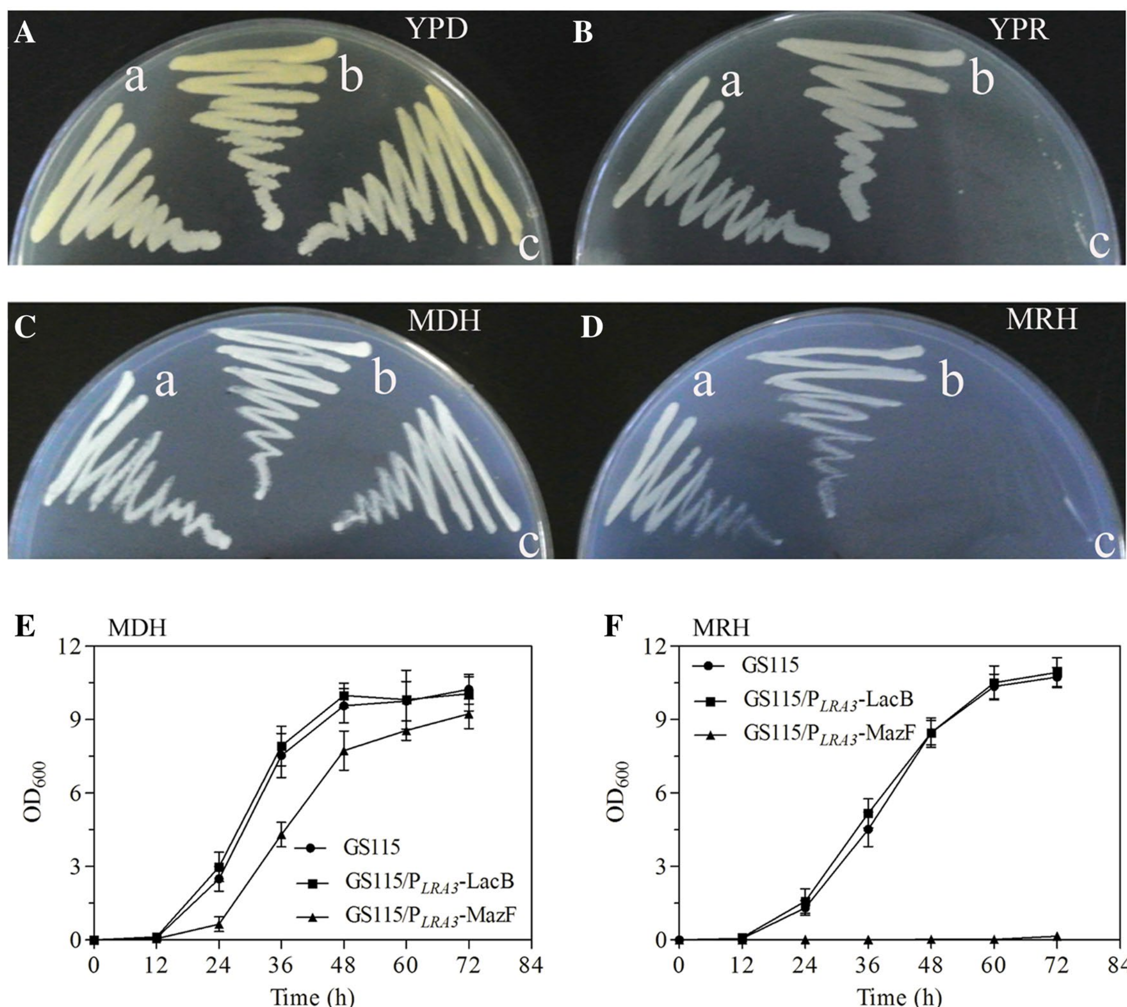

Fig. 2 Growth profiles of *P. pastoris* GS115/P_{LRA3}-MazF grown on/in YPD (**a**), YPR (**b**), MDH (**c**), MRH (**d**), liquid MDH (**e**), liquid MRH (**f**). *P. pastoris* GS115 and *P. pastoris* GS115/P_{LRA3}-LacB were used as the control strains. a, b and c in A, B, C, and D represented the strains *P. pastoris* GS115, *P. pastoris* GS115/P_{LRA3}-LacB, and *P. pastoris* GS115/P_{LRA3}-MazF, respectively. Each test was performed in triplicates, and the results are presented as means ± SD of three replicates

induction. The control strains *P. pastoris* GS115 and *P. pastoris* GS115/P_{LRA3}-*lacB* grew normally on all media (Fig. 2b, d).

The growth profiles were further confirmed by determining cell growth in liquid media. In MDH, *P. pastoris* GS115/P_{LRA3}-MazF was able to grow, but at a relatively slower rate than *P. pastoris* GS115 and *P. pastoris* GS115/P_{LRA3}-LacB (Fig. 2e); this indicated that the low production of MazF due to the basal transcription of P_{LRA3} weakly inhibited the growth of *P. pastoris* GS115//P_{LRA3}-MazF. On the other hand, in MRH, *P. pastoris* GS115/P_{LRA3}-*mazF* was completely inhibited by the severe toxicity of massive MazF production arising from rhamnose induction at P_{LRA3} (Fig. 2f). Taken together, these results further indicated that P_{LRA3} exhibited low basal transcription levels in the absence of rhamnose.

Markerless gene deletion system based on P_{LRA3}

Due to the advantages of P_{LRA3}, which include an almost complete attenuation of transcription in the absence of rhamnose and the maintenance of high transcription activity in the presence of rhamnose, it can be a useful tool for genetic manipulation. Thus, a markerless gene deletion system based on P_{LRA3} was developed in the present study (Fig. 3); six steps were involved in the deletion of a single gene. First, homologous fragment A and homologous fragment C were fused via overlap-PCR and two restriction enzyme sites, *Pme* I and *Sna*B I, were introduced. Second, the DNA fragment B, and the expression cassette *mazF–zeocin*, were inserted into the *Sna*B I and *Pme* I restriction enzyme sites, respectively. Third, homologous recombination occurred between the DNA fragment and the chromosomal DNA via

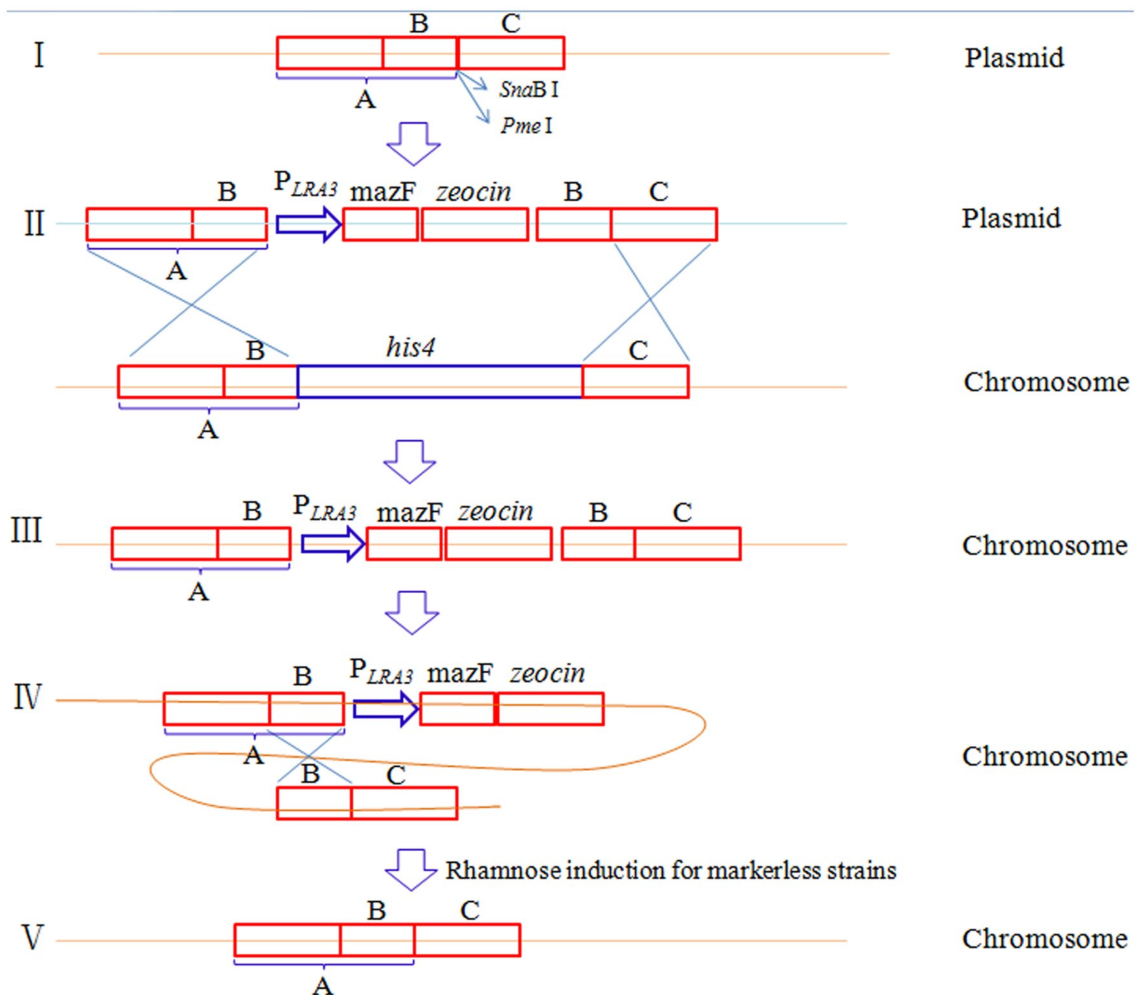

Fig. 3 Schematic diagram of a markerless deletion system in *P. pastoris* GS115

fragment A and fragment C after introduction of the resultant DNA fragment into *P. pastoris* GS115 cells. Fourth, the positive recombinant strain, which was designated as *P. pastoris* GS115-MZ, was isolated with YPDZ and then verified by PCR. Fifth, a second recombination between the two repeats of fragment B led to deletion of the expression cassette *mazF–zeocin*. Cells without the expression cassette *mazF–zeocin* grew normally and were enriched in rhamnose-containing YPR medium.

Markerless gene deletion of *his4*

To evaluate the usefulness and efficiency of this gene deletion system, a *his4* deletion procedure was performed. First, a markerless deletion plasmid, pHISABCZ–*mazF*, was constructed for *his4* and the expression cassette *mazF–zeocin* was integrated into chromosomal DNA at the *his4* locus after its introduction into *P. pastoris* GS115. The positive transformant, which was designated as *P. pastoris* GS115/*his4*Z–*mazF*, was then screened onto YPDZ medium. A single colony of *P. pastoris* GS115/*his4*Z–*mazF* was inoculated into a YPD medium and grown at 28 °C for 48 h while shaking. The cultures were then inoculated into YPD and YPR media at 0.1% (v/v). The OD_{600} values of cultures in the YPD and YPR media were determined simultaneously.

When *P. pastoris* GS115/*his4*Z–*mazF* was grown in YPD, the OD_{600} of the cultures increased rapidly to 9.0 at 36 h (Fig. 4a). Only trace amounts of MazF were produced due to the low basal transcription of P_{LRA3} in YPD and the growth of *P. pastoris* GS115/*his4*Z–*mazF* was slightly inhibited. On the contrary, the OD_{600} of the cultures slowly increased when *P. pastoris* GS115/*his4*Z–*mazF* was grown in YPR and reached 0.213 at 36 h (Fig. 4a). This might have been due to the massive production of MazF following the induction of rhamnose in YPR, which severely inhibited the growth of *P. pastoris* GS115/*his4*Z–*mazF*. Only the cells that abolished the expression cassette *mazF–zeocin* by homologous

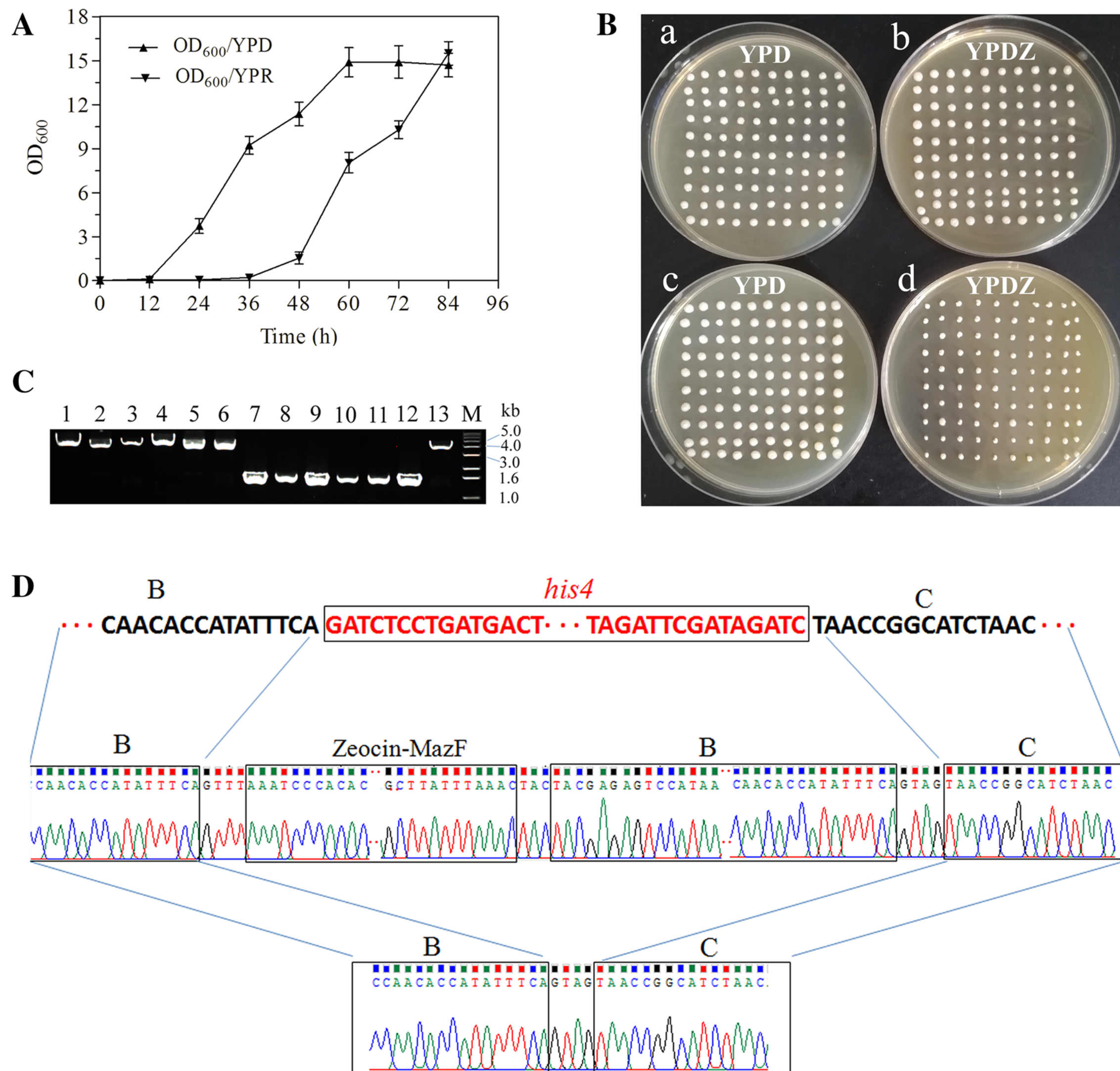

Fig. 4 Markerless deletion of *his4* in *P. pastoris* GS115. **a** Growth profiles of *P. pastoris* GS115/*his4*Z–MazF grown in YPD and YPR. **b** Growth profiles of sub-cultures of *P. pastoris* GS115/*his4*Z–MazF grown in YPD on YPD (**a**) and YPDZ (**b**); growth profiles of sub-cultures of *P. pastoris* GS115/*his4Z–mazF* grown in YPR on YPD (**c**) and YPDZ (**d**). **c** Strain verification via PCR. PCR products for wild-type strain, recombinant strain *P. pastoris* GS115-MZ, and *his4*-deleted strain were ~4.38 kb, ~4.46 kb, and ~1.7 kb, respectively. Lanes 1–6, 7–12, strains from big and small colonies on YPDZ, respectively; lane 13, wild-type strain *P. pastoris* GS115; lane M, DNA marker. **d** Nucleotide assay in *his4*-deleted strain by DNA sequencing

recombination of the two repeats of fragment B could survive and grow in YPR. The OD_{600} of cultures in YPD and YPR at the final incubation of 84 h were similar to previous values.

Theoretically, almost all cells from the cultures of *P. pastoris* GS115/*his4Z–mazF* grown in YPD and YPR should be Zeocin-resistant and Zeocin-sensitive, respectively. As expected, all of the cells collected from the sub-cultures of *P. pastoris* GS115/*his4*Z–*mazF* grown in YPD formed big colonies on YPD and YPDZ (Fig. 4ba, b), which indicates that homologous recombination did not occur between the two repeats of fragment B; thus, the expression cassette *mazF–zeocin* was not deleted. On the other hand, all of the cells collected from the sub-cultures of *P. pastoris* GS115/*his4Z–mazF* inoculated into YPR formed big colonies on YPD (Fig. 4bc) and small colonies on YPDZ (Fig. 4bd). These small colonies on YPDZ might have formed from *mazF–zeocin*-deleted cells and were designated as *P. pastoris* GS115/*his*KO.

To verify these results, DNA fragments containing *his4* were amplified using chromosomal DNA from cells in six big colonies and six small colonies that were randomly selected from YPDZ. The PCR products were of the expected size: 4.38 kb for the wild-type strain, 4.46 kb for

the recombinant strain of *P. pastoris* GS115/*his4*Z–*mazF*, and 1.7 kb for the *P. pastoris* GS115/*his*KO strain (Fig. 4c). The PCR products that responded to the *his4*-deleted strains were sequenced (Fig. 4d), which revealed that *his4* and the expression cassette *mazF–zeocin* were deleted with four nucleotides (GTAG) remaining as the deletion scar. Three nucleotides (GTA) comprised the right part of the recognition sequence 5′-TACGTA-3′ for *Sna*B I, while the fourth nucleotide (G) was introduced after the recognition sequence of *Sna*B I in the primers (*his4A*-R and *his4C*-F). In fact, the fourth nucleotide (G) was not necessary and could not be introduced into the primers; i.e., this system exhibited high efficiency in realizing markerless gene deletion as well as scarless gene deletion, as only three nucleotides remained.

To further verify the validity of this system, a *leu2* knockout was carried out. The construction of a deletion plasmid pLEU2ABCZ–*mazF* harboring the expression cassette *mazF–zeocin* was similar to that of pHISABCZ–*mazF*. The expression cassette *mazF–zeocin* was introduced into *P. pastoris* GS115/*his*KO and then integrated at the *leu2* locus. The transformants screened onto YPDZ were primarily formed by a single step of double-crossover recombination via the left and right homologous arms of *leu2*. A total of 400–800 transformants per microgram of plasmid occurred and the frequency of positive transformants that could not grow on MDH was relatively high, reaching 84% (42/50; Fig. 5a). The selective markerless strains obtained a frequency of nearly 100% after the positive transformants were grown in YPR, which was verified using PCR. The expected sizes of the PCR products were as follows: ~3.0 kb for the wild-type strain, ~3.8 kb for transformants harboring the expression cassette *mazF–zeocin*, and ~1.6 kb for the *leu2*-deleted strain (Fig. 5b).

Mutation strategy based on P_{LRA3} and PB

Transposon insertional mutagenesis has been widely applied to investigate gene function and screen for expected mutants (Barquist et al. 2013). Recently, it was reported that the *piggyBac* transposon under control of the repressible *MET3* promoter is effective in *P. pastoris* (Zhu et al. 2018). Compared to repressible promoters, inducible promoters are advantageous in that they promote the precise expression of a target gene. Thus, in the present study, P_{LRA3} was applied to regulate *piggyBac* expression for the construction of a mutation library of *P. pastoris* (Fig. 6a).

For the mutant library construction, two nucleotides (G and C), an exogenous DNA fragment containing the

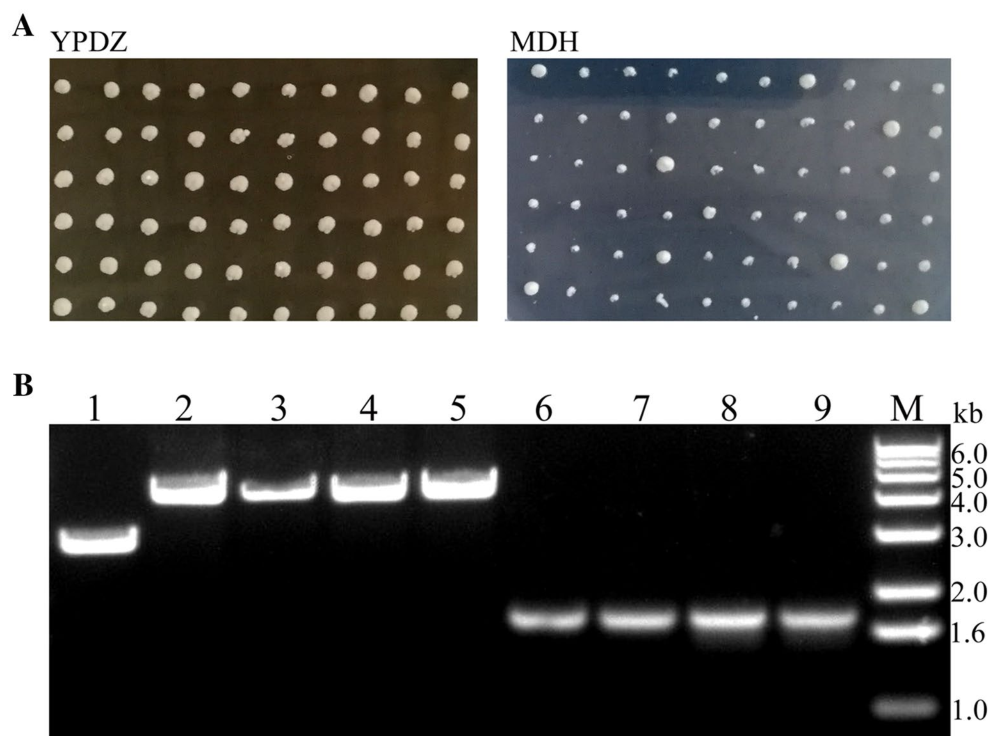

Fig. 5 Markerless deletion of *leu2* in *P. pastoris* GS115/*his*KO. **a** Growth profiles of sub-cultures of transformants grown on YPDZ and MDH. Positive transformants could grew on YPDZ instead of MDH due to *leu2* disruption. **b** Strain verification by PCR. The PCR products were ~3.0 kb for the wild-type strain, ~3.8 kb for transformants harboring the expression cassette *mazF–zeocin*, and ~1.6 kb for the *leu2*-deleted strain

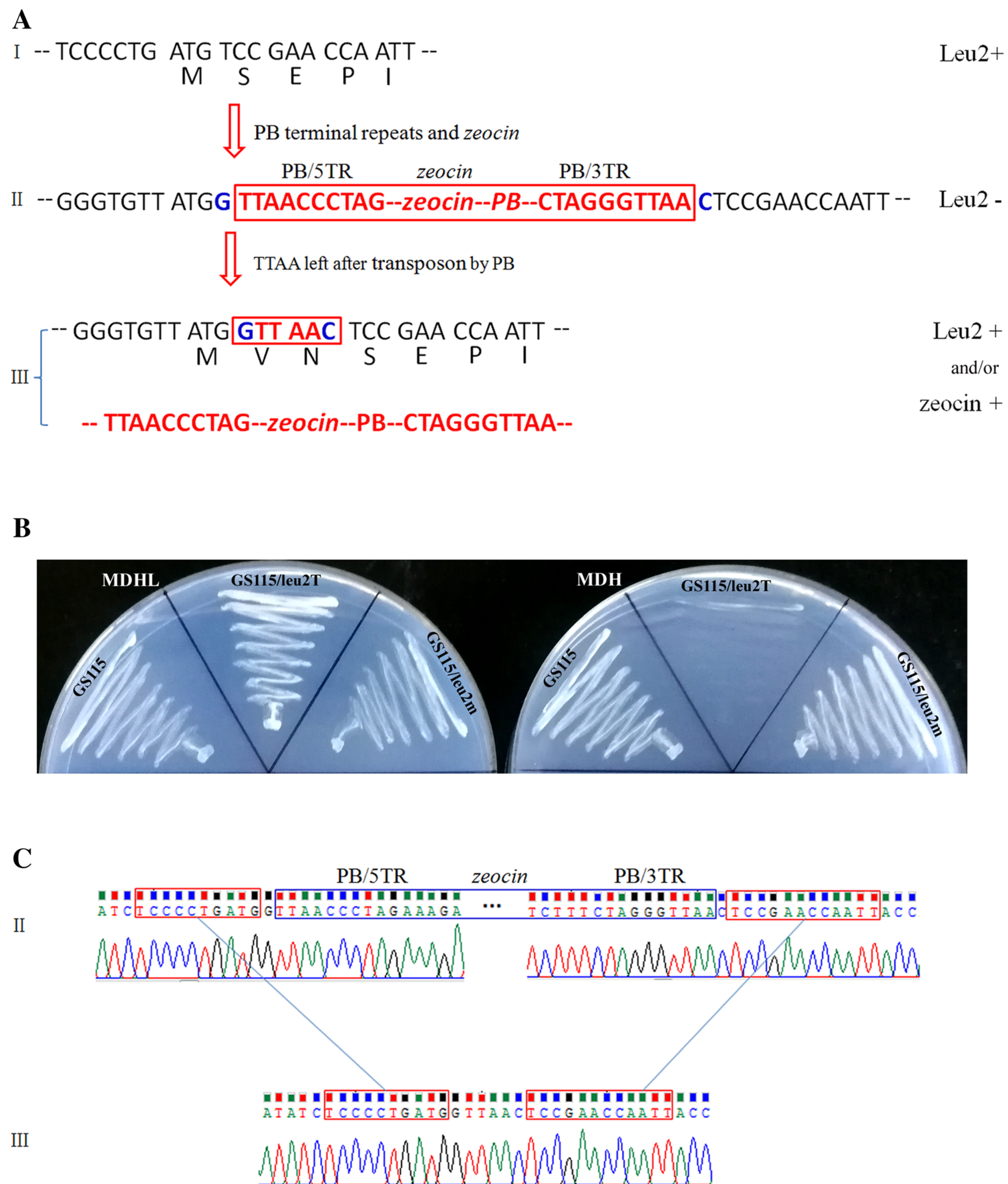

Fig. 6 Mutation system based on P_{LRA3} regulating *PB* expression. **a** Schematic diagram of mutation system in *P. pastoris*. **b** Growth profiles of different strains on media MDHL and MDH. GS115, *P. pastoris* GS115; GS115/leu2T, *P. pastoris* GS115/*leu2T*; GS115/leu2m, *P. pastoris* GS115/*leu2m*. **c** Nucleotide assay of *leu2* locus in *P. pastoris* GS115/*leu2m* by DNA sequencing

PB-recognizing sequence (PB/5TR and PB/3TR), a selective marker (*zeocin*), and *piggyBac* (the expression of which was under the control of P_{LRA3}) were introduced into the CDS of *leu2* via overlap-extension PCR to build the DNA fragment *leu2T*. The *leu2T* fragment was then introduced into *P. pastoris* GS115 and integrated into the *leu2* locus via homologous recombination. Due to the *leu2* disruption, *P. pastoris* GS115 cells harboring *leu2T*, which were designated as *P. pastoris* GS115/*leu2T*, could survive on a leucine-supplemented MDHL medium instead of an MDH medium. When *P. pastoris* GS115/*leu2T* was grown in rhamnose-containing media, such as MRH and YPR, *piggyBac* was intensively expressed. The exogenous DNA fragment in *P. pastoris* GS115/*leu2T* could be excised

from the CDS of *leu2* with four nucleotides (TTAA) remaining. As a result, six nucleotides (GTTAAC) were introduced into the CDS of *leu2* to generate the modified version of *leu2*, which was termed *leu2m*. The *leu2m* gene could encode functional proteins involved in leucine biosynthesis and the *P. pastoris* GS115/*leu2m* strain could survive in the MDH medium. As expected, the *P. pastoris* GS115/*leu2T* strain could not grow on MDH medium but did grow on MDHL medium, whereas the *P. pastoris* GS115 and *P. pastoris* GS115/*leu2m* strains grew well on both media (Fig. 6b). DNA sequencing analyses further verified that the exogenous DNA fragment in *P. pastoris* GS115/*leu2T* was precisely excised from the *leu2* locus with two artificially introduced nucleotides in primers (G and C) and four nucleotides (TTAA) remaining (Fig. 6c).

When the *P. pastoris* GS115/*leu2T* strain was cultured in a YPR or MRH medium, the exogenous DNA fragment could be excised from the *leu2* locus with two resulting possibilities for the exogenous DNA fragment: reinsertion into a new genomic locus of TTAA or loss during/after cell division. The mutation occurred only when the exogenous DNA fragment was reinserted into another locus of TTAA and the mutants could be screened onto MDH medium with Zeocin (100 μg/ml). To determine the frequency of excision and reinsertion of the exogenous DNA fragment from the *leu2* locus when *P. pastoris* GS115/*leu2T* was cultured in MRH, the colony-forming units (CFU) on the YPD, MDH, and MDHZ were assessed. The colonies on YPD were formed from all living cells, the colonies on MDH were formed from all cells in which the exogenous DNA fragment was excised from *leu2*, and the colonies on MDHZ were formed from cells in which the exogenous DNA fragment was excised from *leu2* and then reinserted into another locus of TTAA. Based on the CFU number (Fig. 7), the excision frequency of the exogenous DNA fragment from the *leu2* locus was 61.3% and the reinsertion frequency of the exogenous DNA fragment was 12.01%. These results indicate that this system was highly efficient for producing mutations in *Pichia* and may be an excellent tool to establish a mutant library.

Screening of rhamnose repression-resistant mutants

As a carbon source, rhamnose is preferred over methanol in *P. pastoris*, and methanol utilization is repressed in the presence of rhamnose. P_{AOX1} was not activated in rhamnose-containing media, such as YPR and MR, and *P. pastoris* GS115/P_{AOX1}-GFP did not express GFP using rhamnose as a sole carbon source. To investigate mechanism of rhamnose repression, it was essential to isolate rhamnose repression-resistant mutants. To isolate these mutants, *P. pastoris* GS115/*leu2T*/P_{AOX1}-GFP was thereby developed from *P. pastoris* GS115/P_{AOX1}-GFP. When *P. pastoris* GS115/*leu2T*/P_{AOX1}-GFP was grown in MRHZ medium, mutants occurred due to excision and subsequent insertion into other TTAA locus of the exogenous DNA fragment. The mutants expressed GFP which was under control of P_{AOX1} were considered as rhamnose repression-resistant mutants. Analyzing genomes of these mutants would provided some insights into rhamnose repression and into designing *P. pastoris* GS115 with rhamnose or other carbon sources instead of flammable and hazardous methanol as the inducers to activate P_{AOX1}.

~1,000,000 cells from each culture of *P. pastoris* GS115/P_{AOX1}-GFP grown in MRH, *P. pastoris* GS115/*leu2T*/P_{AOX1}-GFP grown in MRHZ, and *P. pastoris* GS115/P_{GAP}-GFP grown in MRH were analyzed. *P. pastoris* GS115/P_{AOX1}-GFP was used as the negative control, and the gate without green fluorescence signal was drawn as P1 gate. Similarly, *P. pastoris* GS115/P_{GAP}-GFP was used as the positive control, the area with stronger green fluorescence signal as P2 gate. Only two cells from *P. pastoris* GS115/P_{AOX1}-GFP cultures fell into P1 gate, and no cells fell into P2 gate (Fig. 8a). As for *P. pastoris* GS115/P_{GAP}-GFP, 92.2% and 30.9% of cells were in P1 and P2 gates, respectively (Fig. 8b). When

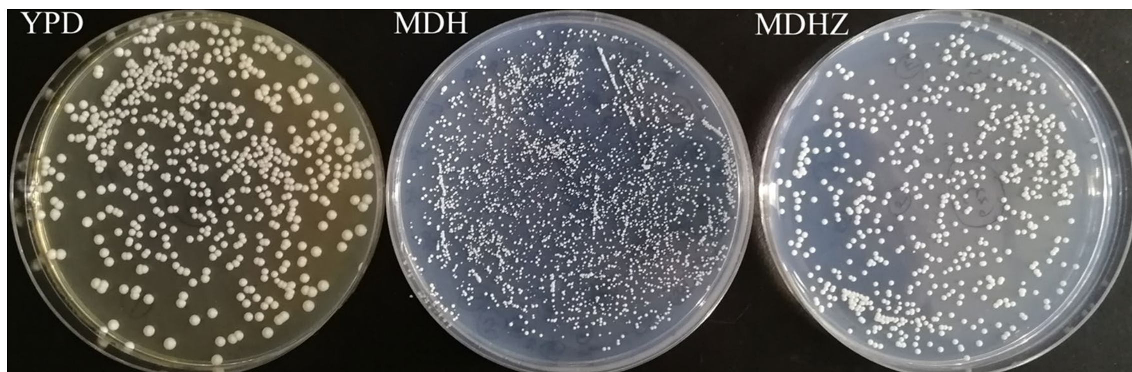

Fig. 7 CFU on different media when *P. pastoris* GS115/*leu2T* was cultured in MRH. 100 μl of cultures after dilution with 10^5-fold, 10^4-fold and 10^4-fold, respectively, were inoculated on YPD, MDH and MDHZ

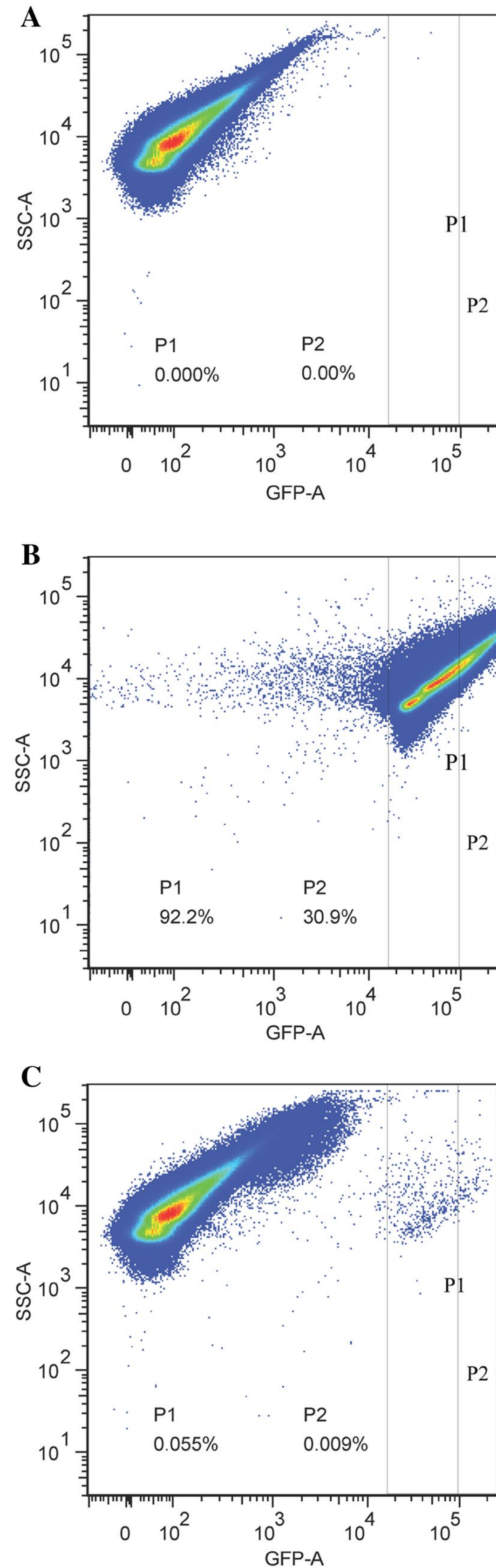

Fig. 8 Green fluorescent intensity of ~1,000,000 cells from each cultures of the following strains. **a** *P. pastoris* GS115/P_{AOX1}-GFP, **b** *P. pastoris* GS115/P_{GAP}-GFP, **c** *P. pastoris* GS115/*leu2T*/P_{AOX1}-GFP. Percentages of cells falling in each gate were indicated

P. pastoris GS115/*leu2T*/P_{AOX1}-GFP was grown in selective medium MRHZ, only mutants could survive. Thus, ~1,000,000 cells from *P. pastoris* GS115/*leu2T*/P_{AOX1}-GFP cultures were the mutant cells. It was shown that 0.055% and 0.009% of these mutant cells occurred in P1 and P2 gates, respectively (Fig. 8c). To avoid false positive samples, cells in P2 gate rather than P1 were considered as the positive mutants, the rhamnose repression-resistant mutants. These mutants were sorted, and some would be verified by genome sequencing to identify the mutation sites. The results further indicated that the developed mutation strategies were highly efficient to screen expected mutants and to investigate functions of specific genes.

Discussion

Inducible promoters are useful tools for controlling the expression of target genes; moreover, tightly regulated promoters are critical to achieve a massive yield of recombinant proteins and stability in engineered strains (McCutcheon et al. 2018; Rosano and Ceccarelli 2014). However, many inducible promoters are disadvantageous due to high basal expression levels; as a result, a variety of strategies have been adopted to reduce basal transcription in some promoters. For example, a T7 lysozyme or a l*acO* operator downstream of T7 promoter can be introduced in *E. coli* to reduce the basal transcription of some promoters (Gopal and Kumar 2013; McCutcheon et al. 2018).

The rhamnose-inducible promoter from *E. coli* is an ideal inducible promoter for the production of recombinant proteins due to its strong transcription as well as for theoretical studies due to its low basal expression. P_{LRA3} from *P. pastoris* also has a great deal of potential for massive production of recombinant proteins (Liu et al. 2016; Yan et al. 2018); thus, the present study conducted two experiments to determine its basal transcription level. The results indicated that P_{LRA3} maintained a low basal transcription level under non-inducible conditions based on two indexes: (1) low β-galactosidase production in the fermentation supernatants of *P. pastoris* GS115/P_{LRA3}-LacB grown in non-inducible media and (2) the slightly slower growth of *P. pastoris* GS115/P_{LRA3}-MazF grown in non-inducible media compared to wild-type *P. pastoris* GS115. However, it was also noted that P_{LRA3} was less tightly regulated than the well-known methanol-inducible promoter P_{AOX1}. Thus, other strategies, such as the introduction of

a glucose repression-related repressor-binding nucleotide sequence into P_{LRA3}, were implemented to reduce leaky expression in another experiment.

Previously, a scarless gene deletion system was developed in *Hansenula polymorpha* using the split-marker method (Song et al. 2015). In this method, transformants with Zeocin resistance occurred via three-crossover homologous recombination, with two being due to homologous recombination between the left/right homologous arms and chromosomal DNA, and the third to the two split markers. Thus, transformation efficiency was low compared to classical transformation using an intact disruption cassette. Using the split-marker method, about 167 colonies of transformants per microgram of DNA occur and half of those are obtained from homologous recombination. In the present study, an intact disruption cassette with double-crossover homologous fragments was constructed in vitro via overlap-extension PCR and T4-ligase-mediated ligation, and then introduced into *P. pastoris* GS115. Transformants occurred following only one step of double-crossover homologous recombination, and about 400–800 colonies of transformants per microgram of DNA occurred, with over 80% of the transformants being positive. Compared to the split-marker method, the present method was advantageous in terms of transformation efficiency and a positive transformant ratio. However, one disadvantage to the present method was a small scar left in the chromosomal DNA via three or four nucleotides. Utilizing this system, a single strain with deletions of *his4, leu2*, and *sdh* plus *gas1* was constructed by subsequent gene deletion, and this strain would be used as the host for expression of multiple secondary metabolite genes to produce expected metabolites in *Pichia*.

Zhu et al. (2018) recently used the *MET3* promoter to regulate *piggyBac* expression and established *piggyBac* transposon-mediated mutagenesis in *Pichia*; the highest transposition frequency using the *MET3* promoter to drive PB expression was 3%. In the present study, a *Pichia* mutation system based on the P_{LRA3}-regulated expression of *piggyBac* was developed and the transposition frequency reached 12%, which was higher than that reported in *Schizosaccharomyces pombe* by Li et al. (2011). These findings indicate that this system could be used to isolate expected mutants and to identify the functions of unknown genes in *Pichia*.

The scarless gene deletion system in combination with the mutation system would be an excellent strategy to develop perfect *Pichia* hosts. Specific genes related to unfavorable profiles was first screened via the mutation system, and then were deleted using the markerless gene deletion system.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant number: 31671802) and Fundamental Research Funds for Central Non-profit Scientific Institution (Grant nos.: 1610392018006, Y2018LM02).

References

Ahn J, Park KM, Lee H, Son YJ, Choi ES (2013) *GAL* promoter-driven heterologous gene expression in *Saccharomyces cerevisiae* delta strain at anaerobic alcoholic fermentation. FEMS Yeast Res 13:140–142

Barquist L, Boinett CJ, Cain AK (2013) Approaches to querying bacterial genomes with transposon-insertion sequencing. RNA Biol 10:1161–1169

Berghoff BA, Wagner EGH (2017) RNA-based regulation in type I toxin–antitoxin systems and its implication for bacterial persistence. Curr Genet 63:1011–1016

Bijlani S, Nahar AS, Ganesan K (2018) Improved Tet-On and Tet-Off systems for tetracycline-regulated expression of genes in *Candida*. Curr Genet 64:303–316

Da-Silva NA, Srikrishnan S (2012) Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. FEMS Yeast Res 12:197–214

Giacalone MJ, Gentile AM, Lovitt BT, Berkley NL, Gunderson CW, Surber MW (2006) Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. Biotechniques 40:355–364

Gopal GJ, Kumar A (2013) Strategies for the production of recombinant protein in *Escherichia coli*. Protein J 32:419–425

Herai S, Hashimoto Y, Higashibata H, Maseda H, Ikeda H, Omura S, Kobayashi M (2004) Hyper-inducible expression system for *streptomycetes*. Proc Natl Acad Sci 101:14031–14035

Li J, Zhang JM, Li X, Suo F, Zhang MJ, Hou WR, Han JH, Du LL (2011) A piggyBac transposon-based mutagenesis system for the fission yeast *Schizosaccharomyces pombe*. Nucleic Acids Res 39:e40

Lim B, Zimmermann M, Barry NA, Goodman AL (2017) Engineered regulatory systems modulate gene expression of human commensals in the gut. Cell 169:547–558

Liu B, Zhang YW, Zhang X, Yan CL, Zhang YH, Xu XX, Zhang W (2016) Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. Sci Rep 6:27352

Luo X, Yang YW, Ling W, Zhuang H, Li Q, Shang GD (2016) *Pseudomonas putida* KT2440 markerless gene deletion using a combination of lambda Red recombineering and *Cre*/*loxP* site-specific recombination. Fems Microbiol Lett 363:fnw014

McCutcheon SR, Chiu KL, Lewis DD, Tan C (2018) CRISPR-Cas expands dynamic range of gene expression from T7RNAP promoters. Biotech J 13:1700167

Meisner J, Goldberg JB (2016) The *Escherichia coli rhaSR*-P*rhaBAD* Inducible promoter system allows tightly controlled gene expression over a wide range in *Pseudomonas aeruginosa*. Appl Environ Microbiol 82:6715–6727

Rodriguez-Garcia A, Combes P, Perez-Redondo R, Smith MCA, Smith MCM (2005) Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. Nucleic Acids Res 33:e87

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:172

Saida F, Uzan M, Odaert B, Bontems F (2006) Expression of highly toxic genes in *E. coli*: Special strategies and genetic tools. Curr Protein Pept Sci 7:47–56

Song P, Liu S, Guo XN, Bai X, He XP, Zhang BR (2015) Scarless gene deletion in methylotrophic *Hansenula polymorpha* by using *mazF* as counter-selectable marker. Anal Biochem 468:66–74

Trassaert M, Vandermies M, Carly F, Denies O, Thomas S, Fickers P, Nicaud JM (2017) New inducible promoter for gene expression and synthetic biology in *Yarrowia lipolytica*. Microb Cell Fact 16:141

Unver Y, Yildiz M, Kilic D, Taskin M, Firat A, Askin H (2018) Efficient expression of recombinant human telomerase inhibitor 1 (hPinX1) in *Pichia pastoris*. Prep Biochem Biotechnol 29:1–6

Yan CL, Xu XX, Zhang X, Zhang YW, Zhang YH, Zhang ZF, Zhang W, Liu B (2018) Decreased rhamnose metabolic flux improved production of target proteins and cell flocculation in *Pichia pastoris*. Front Microbiol 9:1771

Yang J, Cai H, Liu J, Zeng M, Chen J, Cheng Q, Zhang L (2018) Controlling AOX1 promoter strength in *Pichia pastoris* by manipulating poly (dA:dT) tracts. Sci Rep 8:1401

Yu BJ, Kang KH, Lee JH, Sung BH, Kim MS, Kim SC (2008) Rapid and efficient construction of markerless deletions in the *Escherichia coli* genome. Nucleic Acids Res 36:e84

Zhao T, Li Z, Guo Z, Wang A, Liu Z, Zhao Q, Li Y, McKenzie EA, Diao A (2018) Functional recombinant human legumain protein expression in *Pichia pastoris* to enable screening for legumain small molecule inhibitors. Protein Expr Purif 150:12–16

Zhu J, Zhu Q, Gong R, Xu Q, Cai M, Jiang T, Zhou X, Zhou M, Zhang Y (2018) *PiggyBac* transposon-mediated mutagenesis and application in yeast *Komagataella phaffii*. Biotechnol Lett 40:1365–1376

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.