13000–13012 *Nucleic Acids Research, 2020, Vol. 48, No. 22* *Published online 1 December 2020*
doi: 10.1093/nar/gkaa1066

Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression

Yoichiro Ito[1,2], Goro Terai[3], Misa Ishigami[4], Noriko Hashiba[4], Yasuyuki Nakamura[1,2], Takahiro Bamba[1], Ryota Kumokita[1], Tomohisa Hasunuma[1,2], Kiyoshi Asai[3], Jun Ishii[1,2,*] and Akihiko Kondo[1,2,5,*]

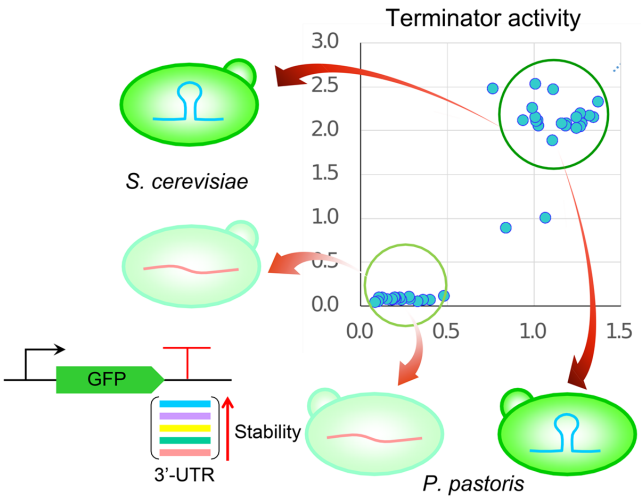
[1]Graduate School of Science, Technology and Innovation, Kobe University, Kobe 657-8501, Japan, [2]Engineering Biology Research Center, Kobe University, Kobe 657-8501, Japan, [3]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, University of Tokyo, Chiba 277-8561, Japan, [4]Technology Research Association of Highly Efficient Gene Design, Kobe 650-0047, Japan and [5]Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe University, Kobe 657-8501, Japan

Received April 28, 2020; Revised October 15, 2020; Editorial Decision October 19, 2020; Accepted October 22, 2020

ABSTRACT

In the yeast *Saccharomyces cerevisiae*, terminator sequences not only terminate transcription but also affect expression levels of the protein-encoded upstream of the terminator. The non-conventional yeast *Pichia pastoris* (syn. *Komagataella phaffii*) has frequently been used as a platform for metabolic engineering but knowledge regarding *P. pastoris* terminators is limited. To explore terminator sequences available to tune protein expression levels in *P. pastoris*, we created a 'terminator catalog' by testing 72 sequences, including terminators from *S. cerevisiae* or *P. pastoris* and synthetic terminators. Altogether, we found that the terminators have a tunable range of 17-fold. We also found that *S. cerevisiae* terminator sequences maintain function when transferred to *P. pastoris*. Successful tuning of protein expression levels was shown not only for the reporter gene used to define the catalog but also using betaxanthin production as an example application in pathway flux regulation. Moreover, we found experimental evidence that protein expression levels result from mRNA abundance and *in silico* evidence that levels reflect the stability of mRNA 3′-UTR secondary structure. In combination with promoter selection, the novel terminator catalog constitutes a basic toolbox for tuning protein expression levels in metabolic engineering and synthetic biology in *P. pastoris*.

GRAPHICAL ABSTRACT



INTRODUCTION

Numerous valuable commodities have been produced in yeasts using metabolic engineering technology. Design of an optimal yeast strain for production of a target compound often requires regulation of gene expression, including overexpression, fine-tuned expression, knock-down and knockout, which together can be used to optimize pathway flux for production of the target and can rewire cellular metabolism (1,2). To tune gene expression for pathway optimization, researchers have mainly explored modifications to 5′-untranslated regions (5′-UTRs) (3–7). For example, we have used promoter catalogs resulting from many studies to

*To whom correspondence should be addressed. Tel: +81 78 803 6356; Fax: +81 78 803 6192; Email: junjun@port.kobe-u.ac.jp
Correspondence may also be addressed to Akihiko Kondo. Tel: +81 78 803 6196; Fax: +81 78 803 6196; Email: akondo@kobe-u.ac.jp

select promoters with the desired level of activity, e.g. high, medium, or low (6–10). Although this type of promoter-exchange strategy can result in a wide tunable range of protein expression levels, it is complicated by the observation that changes in culture conditions, such as carbon sources or stressors, might influence promoter activity (10–12) and thus, affect production of the target materials.

Another important tool used to regulate protein expression levels is terminator exchange. Recently, it was reported that terminator regions not only function in transcriptional termination but can also influence expression levels of a protein encoded upstream of the terminator (13,14). In general, compared with promoter studies, few studies of terminators useful to tune protein expression of a gene of interest have been reported (15). For *Saccharomyces cerevisiae*, however, in one large-scale study, a catalog of 5302 terminators was developed and characterized (13). The terminators resulting in the highest and lowest expression levels were *DIT1*t and *GIC1*t, respectively, which differed by more than 70-fold (13,16). These results demonstrate that by choosing a terminator sequence, protein expression levels can be regulated within a broad range. Moreover, the *S. cerevisiae* endogenous terminators were shown to affect protein expression independently of the upstream promoter and reporter open reading frame (ORF) (13). Studies have shown that in *S. cerevisiae*, terminator choice influences mRNA abundance (13,16), mRNA stability (14) and post-transcriptional interaction with endogenous *trans*-acting factors (17). The *trans*-acting factors Nab6p and Pap1p, which were identified in a genome library screen, appear to be responsible for the strong activity of the endogenous *DIT1* terminator in *S. cerevisiae* (17). Artificial terminators have been developed in *S. cerevisiae* using either of two strategies: connection of three functional parts of *S. cerevisiae* terminators (18,19) or insertion of additional *trans*-acting factor binding sites of the *DIT1* terminator into an endogenous terminator sequence (17). Notably, terminator technology based on use of native or artificial terminators has been successfully applied in metabolic engineering (15,19).

The methylotrophic yeast *Komagataella phaffii*, commonly known as *Pichia pastoris*, is widely used as a platform for production of heterologous proteins (20,21). This yeast has also attracted attention as a potential platform for bulk chemical production via metabolic engineering and synthetic biological technologies (22–24). Application of a terminator-exchange strategy has promise for improving production of valuable commodities by *P. pastoris*. However, only a limited number of terminators have been established for use in *P. pastoris*: fifteen endogenous terminators from methanol assimilation pathway genes and five *S. cerevisiae* terminators (10). These 20 terminators showed comparable activities, with a difference among them of only about 2-fold—i.e. terminator activities ranging from 57% (*Pp_CAT1*t) to 100% (*Pp_AOX1*t)—and thus provide only a limited ability to regulate pathway flux.

Terminator sequences consist of a 3′-untranslated region (3′-UTR) followed by a region needed to terminate transcription of the mRNA. Previous studies of eukaryotic 3′-UTRs have reported many sequence motifs within terminators that might affect mRNA stability (25,26). Some of these have been experimentally verified. For example, the AU-rich element is well known to destabilize mRNA in mammals (27). Other studies have reported the contribution of GC content, length and secondary structure of 3′-UTRs to mRNA stability (28–30). By contrast, few studies have addressed the relationship between 3′-UTR sequence and mRNA abundance in *P. pastoris*.

In the current study, we created a catalog of 72 terminators for use in *P. pastoris* that includes endogenous, heterogeneous, and synthetic terminators, and can be used to tune expression levels of a target protein via a terminator-exchange strategy. Using these terminators, we presented the first demonstration of the utility of a terminator-exchange strategy to regulate metabolic pathway flux in *P. pastoris*. Furthermore, we found that protein expression levels inversely correlate with the folding free energy of the associated 3′UTRs, suggesting that stable secondary structure contributes to higher protein expression levels. Finally, we discuss the transferability of *S. cerevisiae* terminator sequences to other yeast and speculate on the existence of *trans*-acting factors that interact with the terminator sequence to influence activity in *P. pastoris*.

MATERIALS AND METHODS

Strains and media conditions

The *P. pastoris* wild-type strain CBS7435 (NRRL-Y11430) was used in this study. *P. pastoris* strains were grown in YPD medium [10 g/l yeast extract (Nacalai Tesque, Kyoto, Japan), 20 g/l Bacto peptone (BD Biosciences, San Jose, CA, USA) and 20 g/l glucose], BMGY, BMMY or BMDY [10 g/l yeast extract, 20 g/l hipolypeptone (Nihon Pharmaceutical, Tokyo, Japan), 13.4 g/l yeast nitrogen base without amino acids (BD Biosciences), 0.4 mg/l biotin (Nacalai Tesque), 100 mM potassium phosphate buffer (final, pH 6.0) and 20 g/l glycerol (for BMGY), 20 g/l methanol (for BMMY) or 20 g/l glucose (for BMDY)]. *Escherichia coli* strains were grown in LB medium that contained: 5 g/l yeast extract, 10 g/l tryptone (Nacalai Tesque) and 5 g/l NaCl, supplemented with ampicillin (100 μg/ml). YPD agar plate contained: YPD medium, 20 g/l agar and antibiotics (500 μg/ml G418; Wako Pure Chemical Industries, Osaka, Japan) and 100 μg/ml hygromycin (Wako Pure Chemical Industries).

The *E. coli* strain DH5α was used for recombinant DNA manipulation. *Escherichia coli* strains were grown in LB medium that contained (per liter) 5 g yeast extract, 10 g tryptone (Nacalai Tesque) and 5 g NaCl, and was supplemented with ampicillin (100 μg/ml).

Construction of plasmids and *P. pastoris* strains

Standard recombinant DNA manipulation was performed as described by Sambrook *et al.* (31). The primers and synthetic DNA fragments used in this study are listed in Supplementary Tables S1 and 2, respectively. Primers were purchased from Eurofin Genomics K. K. (Tokyo, Japan) and synthetic DNA fragments were synthesized by Geneart Strings DNA Fragments (Thermo Fisher Scientific, Waltham, MA, USA) with codon usage optimized for *P. pastoris*.

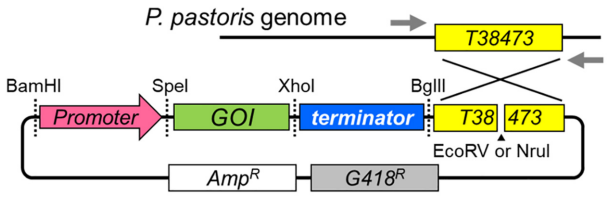


Figure 1. Genome-integrated gene constructs for *Pichia pastoris*. *P. pastoris, Saccharomyces cerevisiae* or synthetic terminators were inserted as modules downstream of a promoter (*GAPDH* or *AOX1* promoter) and a gene of interest (GOI): EGFP, E2Crimson, *Rhizopus oryzae* glucoamylase secretion signal-BGL1p and tyrosine hydroxylase (CYP76AD1)-coding sequences. Following linearization of the plasmid with EcoRV or NruI, homologous integration occurred in the terminator region of *CCA38473* (*T38473*). A G418 resistance-encoding marker was used for selection. Gray arrows indicate primer positions used to identify correct, single integrations by colony PCR.

All terminator plasmids used in this study were based on the plasmid pPGP_EGFP (32), which consists of the *GAPDH* promoter, an EGFP-encoding ORF and an *AOX1* terminator (Figure 1). Detailed methods for construction of the plasmids used in this study are described in the Supplementary Data.

All of the terminator plasmids were linearized by EcoRV or NruI, followed by integration into the CCA38473 terminator locus on the *P. pastoris* genome by use of a LiOAc transformation method, as described in our previous paper (32). Single copy integration of each linearized plasmid into the correct locus was verified by quantitative polymerase chain reaction (qPCR) and colony PCR. The qPCR analysis was performed as previously reported (33). Colony PCR was performed using DNA polymerase (KOD FxNeo, TOYOBO, Osaka, Japan) and primers that recognize sequences upstream and downstream of the integrated loci, according to the manufacturer's protocol.

To optimize DOD expression levels for production of betaxanthin, *P. pastoris* strains carrying various copy numbers of DOD expression cassettes were prepared. The plasmid pPGPH_DOD was linearized by BsrGI and then electroporated into the *ARG4* locus on the *P. pastoris* genome as described in our previous paper (32). Transformants were obtained on YPD plates supplemented with 100 μg/ml hygromycin. The copy number of the DOD expression cassette in each transformant was measured by qPCR as previously reported (33).

Cultivation in test tubes

To quantify protein expression levels in the terminator strains, test tube cultivation was performed as follows. At least three biologically distinct colonies of each yeast strain were streaked on YPD plates supplemented with G418, then were inoculated into 2 ml of BMGY medium in test tubes, followed by pre-culture at 30°C for 24 h at 170 rpm. Next, 200 μl of the pre-culture was transferred into 2 ml of BMMY medium (for *AOX1* promoter strains) or BMGY medium (for *GAPDH* promoter strains) in test tubes, and further cultured at 30°C for 24 h or 48 h at 170 rpm for protein production. After cultivation, biomass was measured by OD_{660} and protein expression levels of target protein levels were determined as described below.

Flow-cytometry analysis and data processing

Transformants producing EGFP (32) or the red fluorescent protein E2Crimson (RFP, (34)) were cultured in test tubes as described above. Fluorescent intensities of 10 000 cells of each terminator strain were measured using a flow-cytometer (FCM, FACS CantoII, BD Bioscience, Franklin Lakes, NJ, USA) equipped with 488- and 633-nm lasers and the appropriate filter set for EGFP (530/30-nm) and RFP (660/20-nm). Data processing was performed as follows: each cell fluorescent intensity (FI) value for EGFP or RFP was divided by the forward scatter intensity to account for variability in cell size, then a histogram was created from the calculated values, fitted by a normal distribution. Terminator activity was defined as the median value of the normal distribution for a given cell sample.

β-glucosidase activity assay

Transformants producing *Aspergillus aculeatus* β-glucosidase (BGL1p) were cultured in test tubes as described above. After centrifugation of the culture media, BGL1p activity in the yeast culture supernatant was assayed with p-nitrophenyl-β-D-glucopyranoside (pNPG) as the substrate as previously reported (35). In brief, 2 μl the culture supernatant was added to 98 μl of pNPG solution dissolved in 50 mM sodium citrate buffer, pH 5.0 (final concentration of pNPG: 2 mM). After mixing and incubation at 25°C for 10 min, 100 μl of 3 M sodium carbonate solution was added to terminate the reaction, and the amount of p-nitrophenol generated (yellow color) was determined based on the absorbance at 405 nm using a micro-plate reader (Synergy HTX, BioTek, Winooski, VT, USA).

Evaluation of betaxanthin production

Transformants producing betaxanthin were cultured in test tubes as described above. Test tube culture supernatants (150 μl each) were then transferred to a 96-well black plate and the FI was determined (excitation: 485/20-nm, emission: 528/20-nm, gain: 50) using a microplate reader (Synergy HTX, BioTek) to measure betaxanthin levels.

Determination of transcript levels, mRNA half-life and 3′-UTR length

The mRNA levels of transgenes with different terminators were measured as follows. First, total RNA from each *GAPDH* promoter-EGFP-terminator strain was isolated from a 6 h culture in BMGY medium using an RNeasy Mini kit (Qiagen, Hilden, Germany). Next, cDNA templates were synthesized from 0.5 μg of total RNA using a cDNA synthesis kit (ReverTra Ace qPCR RT Master Mix with gDNA Remover, TOYOBO). To quantify target cDNAs, real-time PCR was performed using a SYBR-based qPCR kit (KOD SYBR qPCR Mix, TOYOBO) and a quantitative PCR system (Mx3005P, Agilent Technologies, Santa

Clara, CA, USA). A DNA fragment of the *ACT1* gene was used as an internal standard.

The length of the 3′-UTR of each terminator was determined by 3′-RACE analysis using a detection kit (3′-RACE Core Set, Takara Bio, Shiga, Japan), following the manufacturer's instructions. Briefly, cDNA templates were synthesized from 1 μg total RNA as described above using a poly-dT primer. The cDNA templates were then amplified by PCR using primers positioned on the EGFP gene and the poly-dT primer. The amplified DNA fragments were cloned using a DNA cloning kit (Zero Blunt TOPO PCR Cloning Kit, Thermo Fisher Scientific). After transformation of DH5α, eight colonies were randomly chosen and the plasmids were extracted. Then, the 3′-UTRs were determined by Sanger sequencing.

The mRNA half-life was measured by the transcriptional shut-off method described in a previous report (36) with some modifications. Briefly, to measure mRNA half-life, we used the EGFP expressing strains carrying the representative terminators under the control of the methanol-inducible *AOX1* promotor as constructed above. Strains were grown in BMMY (methanol) medium to mid-log phase, then re-suspended in an equal amount of BMDY (glucose) medium. After the medium exchange, samples were collected at the following times: 0, 2, 4, 6, 8, 10, 15 and 20 min. Next, cell pellets were collected following a brief centrifugation, dipped quickly into the liquid nitrogen, and stored at −80°C until cell disruption. Cells were suspended in a phenol extraction solution (Isogen II, Nippon Gene, Tokyo, Japan) and disrupted using a Multi-beads shocker (Yasui Kikai, Osaka, Japan) with 0.5 mm glass bead (#G8772, Sigma-Aldrich, St. Louis, MO, USA), and then total RNA was extracted by following with manufacture's instruction. The EGFP mRNA level was determined using the qPCR method described above and mRNA half-lives were then determined as described previously (36).

in silico analysis of the stability of 3′-UTR secondary structure

We used the minimum free energy (MFE) as a measure of the stability of RNA secondary structure in the 3′-UTR. MFEs were calculated using the RNAfold program in the ViennaRNA package (37). We extracted the fixed length sequences downstream of the stop codon and compared their MFE to normalize the effect of sequence length on the MFE. Specifically, with the 3′-UTR length set to L, we extract L nt sequences downstream from the stop codon and calculate their MFE. Then, we calculate the Spearman's rank correlation coefficient (Spearman's rho) between the mRNA level and MFE. Similarly, the Spearman's rho between the mRNA level and GC content was also calculated. For the terminators whose mRNA half-life and GFP expression level (terminator activity) were measured, we also conducted the same calculation of Spearman's rho using the mRNA half-life or terminator activity instead of the mRNA level. Note that for a given L value, the terminators with a 3′-UTR length experimentally identified in this study to be less than L were excluded.

As reported in the Results section, we observed an inverse correlation (negative rho value) between the mRNA level and MFE, as well as between the mRNA half-life or terminator activity and MFE, for various L. To evaluate the statistical significance of the inverse correlation, we conducted the following statistical test. First, we obtained the minimum rho value (rho^{min}) among Spearman's rho values calculated for various L. Next, we obtained the null distribution of rho^{min} based on randomized data. Specifically, we randomly shuffled the relationship between terminators and their mRNA levels, then calculated the minimum rho value ($rho^{min|R}$), using the same way to calculate rho^{min}. We repeated this calculation and obtained 10^5 $rho^{min|R}$ values. Finally, the P value of rho^{min} was defined as $c/10^5$, where c is the number of $rho^{min|R}$ smaller than rho^{min}. The same statistical tests were also conducted using the mRNA half-life or terminator activity instead of the mRNA level.

We also conducted nucleosome occupancy predictions, because a recent study reported a correlation between terminator activity with predicted nucleosome occupancy in 3′-UTR (38). In accordance with a recent study (38), we calculated a nucleosome occupancy score for each 3′-UTR. Briefly, we calculated the probability of nucleosome occupancy at each base using the nucleosome prediction engine NuPoP (39), on a region consisting of 3.0 kb upstream and 2.0 kb downstream of each 3′-UTR. Then, a nucleosome occupancy score for each 3′-UTR was calculated by summing the occupancy probability of L nucleotides downstream of the stop codon. Then, we calculated the Spearman's rho between the mRNA level and nucleosome occupancy score for various L. For the terminators whose mRNA half-life and terminator activity were measured, we also calculated Spearman's rho between the nucleosome score and the mRNA half-life or terminator activity.

RESULTS

Activities of endogenous, heterologous and synthetic terminators in *Pichia pastoris*

To identify terminator sequences with different activities in *P. pastoris*, we constructed yeast strains carrying various terminator sequences downstream of a GFP-encoding ORF under the control of a constitutive *GAPDH* promoter derived from *P. pastoris* (Figure 1), then profiled a terminator catalog based on GFP FI to define terminator activity. To create a 'terminator catalog' for *P. pastoris*, we chose 72 terminator sequences representing the following five categories (see also Table 1): (i) *S. cerevisiae* terminator sequences with the top 21-ranked and bottom 18-ranked activities in *S. cerevisiae* (13); (ii) *P. pastoris* terminator sequences with high amino acid sequence similarity to the ORFs upstream of the 30 top- and bottom-ranked *S. cerevisiae* terminators (Supplementary Table S3) (13); (iii) previously reported *P. pastoris* terminator sequences with high activities (10); (iv) artificial short terminators originally designed for use in *S. cerevisiae* (18,19); and (v) terminator sequences that served in our study as standards. In total, 72 terminator sequences were inserted into the GFP expression plasmid (Figure 1), and each resulting plasmid was linearized and transformed into *P. pastoris*.

The relative activity of each terminator was assessed by determining GFP expression levels. To ensure a fair comparison, it was essential that each of the transgenes be in-

Table 1. Terminator sequences used in this study

Category	Description	Derivation	# of choice	Ref.
i	Endogenous *S. cerevisiae* terminators with the top 21 and bottom 18-ranked activities in *S. cerevisiae*	*S. cerevisiae*	39	12
ii	Endogenous *P. pastoris* terminators selected as they are downstream of ORFs with sequence similarity to the *S. cerevisiae* ORFs associated with the top or bottom 30-ranked terminators in *S. cerevisiae*	*P. pastoris*	22	12
iii	*P. pastoris* terminator sequences with high reported activities in *P. pastoris*	*P. pastoris*	4	9
iv	Synthetic short terminators	-	3	17,18
v	Terminators used as standards	*P.pastoris S. cerevisiae*	4	-

serted at the same locus in the *P. pastoris* genome and at a single copy number. Therefore, only strains in which a single copy transgene had integrated into the target locus in the *P. pastoris* genome were evaluated in this study. After selection of single GFP-expression transgene strains for each construct, at least three independent transformants were subjected to further evaluation. GFP FI of transformant cells was measured by FCM. In the single cell measurement of the GFP expressing cells, we observed that each terminator strain had a large cell-to-cell variability with regards to both GFP FI and forward-scatter intensity (FS), and the FI and FS values of each cell were correlated (Supplementary Figure S1), indicating that the FI value of the cell was size-dependent because the FS value reflects the cell size. Although the averaged FI has often been used as a parameter for GFP expression levels of the yeast strains in FCM analysis, this could potentially be misleading if we want to know the GFP expression level per unit volume of yeast cell. Since determining the GFP expression level independent of cell size is important to compare the terminator activities, we normalized the FI value of the corresponding cell. Thus, normalized GFP FI/FS was used as a measure of terminator activity in this study, consistent with what was done in a previous study (13).

The activities of the 72 terminators are shown in Figure 2 and Supplementary Table S4. As was observed for *S. cerevisiae* (13), we found in *P. pastoris* that the terminator strains showed a variety of activities and that the activities were inherent to a given terminator sequence. The highest performance was obtained with *Pp_AOX1*t and the lowest with *Sc_GIC1*t, with an overall range in activity among the terminators of about 17-fold. The strongest *Pp_AOX1*t we observed corresponds to one of the most often and traditionally used terminators for heterologous protein expression in *P. pastoris*. Interestingly, the *P. pastoris* strain carrying *Pp_AOX1*t was the only strain among the strains harboring the most active 21 endogenous terminators in *S. cerevisiae* (13) in which terminator activity increased over cultivation time in a time-course analysis (Figure 3). The weakest terminator among the 72 tested in this study, *Sc_GIC1*t, was also the weakest in *S. cerevisiae* (13). In this study, we also characterized the *P. pastoris* strains with each of three different synthetic short terminators, T_{Guo}, $T_{synth27}$ (18) and Syn^{Ter10} (19), which had 2.3-, 3.5- and 5.5-fold higher terminator activity, respectively, than *Sc_CYC1*t in *S. cerevisiae*. In *P. pastoris*, the activities of these synthetic terminators were the same or a little higher (1.0- to 1.2-fold) as compared with *Sc_CYC1*t, such that the synthetic terminators are in the middle range of activities observed for the full terminator catalog (Figure 2 and Supplementary Table S4).

For terminator sequences in categories I and II (Table 1), we were able to compare terminator activity in *P. pastoris* to what had been observed previously in *S. cerevisiae* (13). With category I terminators, we found that the activities of endogenous *S. cerevisiae* terminators in *P. pastoris* roughly correlated with those observed in *S. cerevisiae* (Figure 4A, $R^2 = 0.87$). This showed that the *S. cerevisiae* terminator sequences have the same relative activity in *P. pastoris* as in *S. cerevisiae*, suggesting that *S. cerevisiae* terminator sequences have a high degree of transferability to *P. pastoris*. Moreover, this suggests that the terminator activity depends on the nucleotide sequence itself and not the host. In contrast, for category II terminators, i.e. endogenous *P. pastoris* terminators downstream of ORFs orthologous to *S. cerevisiae* ORFs upstream of characterized terminators (Supplementary Table S3), we did not observe a correlation between activity in *P. pastoris* and activity in *S. cerevisiae* (Figure 4B, $R^2 = 0.067$). Taken together, we conclude that terminator activity is influenced by the nucleotide sequence itself, not by the functionality of the protein upstream of the terminator sequence.

Exchange of upstream promoters or reporter genes

To determine whether or not terminator activity *in P. pastoris* is dependent on upstream sequences, a subset of 10 terminators representative of a range of activities were chosen for further study. The terminators included were *Sc_RPL3*t, *Sc_BNA4*t, *Sc_PGK1*t, *Sc_CYC1*t, *Sc_FUR4*t, *Sc_MIG2*t and *Sc_ICY2*t from *S. cerevisiae*, and *Pp_AOX1*t, *Pp_GAPDH*t and *Pp_GAP1*t from *P. pastoris* (Figure 2). Using the 10 terminators, we performed promoter-swap experiments to evaluate if the effects on expression of terminator and promoter sequences are dependent or independent of one another. To test this, we swapped the *GAPDH* promoter of each *EGFP*-terminator construct for the methanol-inducible *AOX1* promoter, linearized the resulting constructs and introduced them into *P. pastoris*. As shown in Figure 5A and Supplementary Table S5, the terminator activities of the *AOX1* promoter strains correlate with those of the *GAPDH* promoter strains ($R^2 = 0.94$), suggesting that terminator activity is independent of the promoter.

Next, a reporter ORF swap experiment was performed using two reporter proteins, an intra-cellular expressing RFP and a secretory protein, BGL1p. To do this, the EGFP-encoding ORF of each terminator construct was swapped

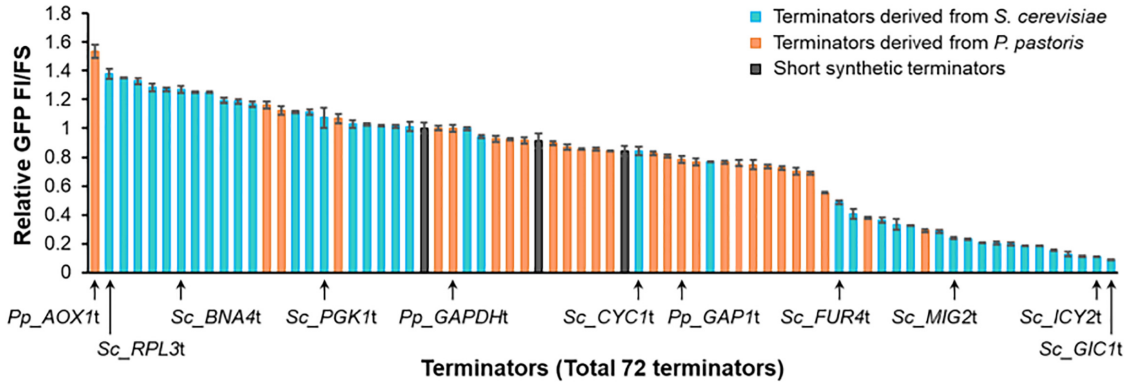


Figure 2. Relative terminator activities of *Pichia pastoris* strains expressing EGFP under the control of the *GAPDH* promoter with various terminators. Blue, activities of the 41 terminators from *Saccharomyces cerevisiae*; orange, the 28 terminators from *P. pastoris*; black, the three synthetic terminators. The terminator activities (relative GFP FI/FS) of all strains are relative to that of the *Pp_GAPDH* terminator strain. The 10 terminators indicated here were used in subsequent experiments. Error bars represent standard deviation from at least three biological replicates.

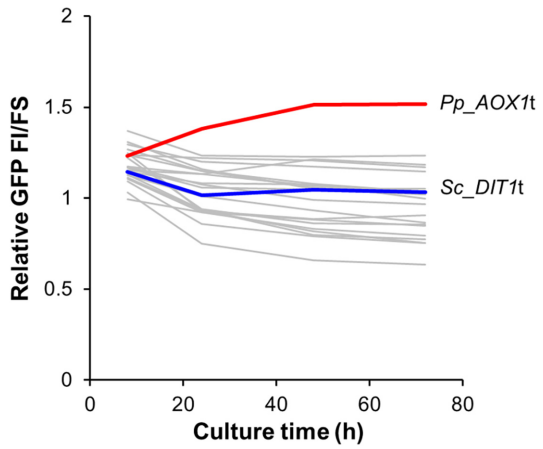


Figure 3. Time-course EGFP expression analysis of terminator activity in *Pichia pastoris*. Shown, the terminator activities (relative GFP FI/FS) of *P. pastoris* strains expressing EGFP under the control of the *GAPDH* promoter with various terminators. The top 21-ranked *Saccharomyces cerevisiae* terminators and *Pp_AOX1* terminator are indicated. Each EGFP FI value relative to that of the *Pp_GAPDH* terminator strain for each sampling time. The terminator activity of the *Pp_AOX1*t *P. pastoris* strain (red line) increased with increased culture time. The *Sc_DIT1*t strain (blue line) showed no apparent increase in terminator activity with increased culture time.

to the RFP coding sequence and BGL1p coding sequence fused with a secretion signal sequence derived from *Rhizopus oryzae* glucoamylase (35). RFP- and BGL1p-producing yeast strains with the 10 terminators were constructed and expressed constitutively under the control of the *GAPDH* promoter. For the RFP constructs, we found that RFP fluorescence in each RFP-terminator strain correlated with GFP fluorescence of the corresponding EGFP-terminator strain (Figure 5B and Supplementary Table S5, $R^2 = 0.97$). For the secretion construct, BGL1p secretion also correlated with GFP fluorescence of the corresponding EGFP-terminator strain (Figure 5C and Supplementary Table S5, $R^2 = 0.96$). Together, these data provide further supporting evidence that terminator activity would be independent of the upstream gene, consistent with what has been reported previously for *S. cerevisiae* (13).

Terminator exchange can be used to tune metabolic pathway flux in *P. pastoris*

As an example application of terminator technology in metabolic engineering, we chose production of the yellow tyrosine-derived pigment betaxanthin. Betaxanthin can be produced in *S. cerevisiae* from L-tyrosine in a three step reaction (40). In step one, L-DOPA is produced from tyrosine using a L-tyrosine hydroxylase mutant strain (CYP76AD1$^{W13L, F309L}$) derived from *Beta vulgaris*. In step two, betalamic acid is produced from L-DOPA via DOPA deoxygenase (DOD) derived from *Mirabilis jalapa*. In the third and final step, betalamic acid reacts spontaneously with amino acids, resulting in production of the yellow fluorescent pigment betaxanthin (Figure 6A). To check whether synthesis of betaxanthin is possible in *P. pastoris*, and furthermore, to modulate betaxanthin metabolic pathway flux in yeast cells, the 10 terminators used in the swap experiments described above were inserted downstream of the double-mutated CYP76AD1 coding sequence under the control of the constitutive *GAPDH* promoter.

To successfully tune metabolic pathway flux for betaxanthin production, it is important to express high levels of DOD, which catalyzes the tyrosine hydroxylase reaction. First, yeast strains carrying various copy numbers of the DOD transgene were prepared by electroporation. We then identified the copy number of DOD transgenes in the genomes of these strains. We finally chose yeast strains carrying one, two or nine copies of the DOD transgene (Supplementary Figure S2). In a preliminary test, CYP76AD1 transgene constructs with *Pp_AOX1* and *Sc_ICY2* terminators were separately transformed into the various DOD copy number strains. After identification of single copy number CYP76AD1 transgene transformants, betaxanthin production levels in each *Pp_AOX1*t and *Sc_ICY2*t strain were evaluated by measuring fluorescence of culture supernatants. For both the *Pp_AOX1*t and *Sc_ICY2*t constructs, betaxanthin levels in the one, two and nine-copy strains were almost the same (Supplementary Figure S2). These

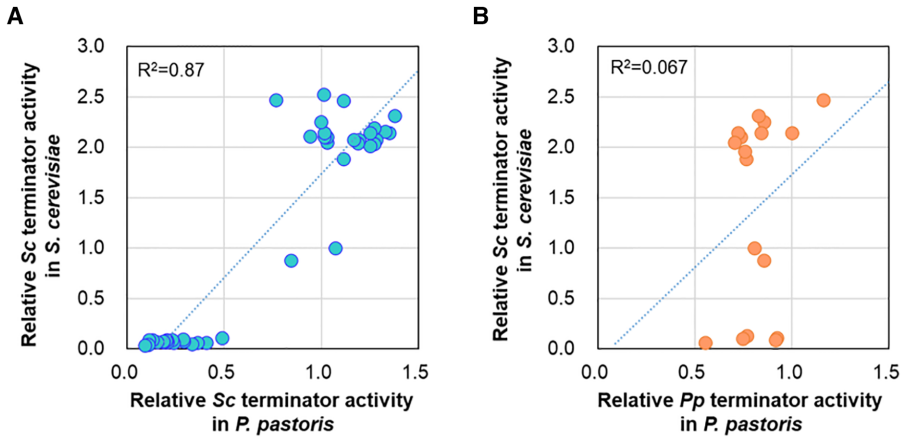


Figure 4. Comparison of the terminator activity between in *Saccharomyces cerevisiae* (vertical axis) and in *Pichia pastoris* (horizontal axis). Terminator sequences in category I (**A**) or II (**B**) are shown. Category I is comprised of endogenous *S. cerevisiae* terminator sequences in the top 21- and bottom 18-rank for activity in *S. cerevisiae*. Category II includes endogenous *P. pastoris* terminator sequences downstream of ORFs with high amino acid similarity to ORFs upstream of *S. cerevisiae* terminators in the top 30- or bottom 30-ranked activities in that species. Data for terminator activities in *S. cerevisiae* are from a previous report (13). Terminator activities (relative GFP FI/FS) in *P. pastoris* were relative to those of the *Pp_GAPDH* terminator strains.

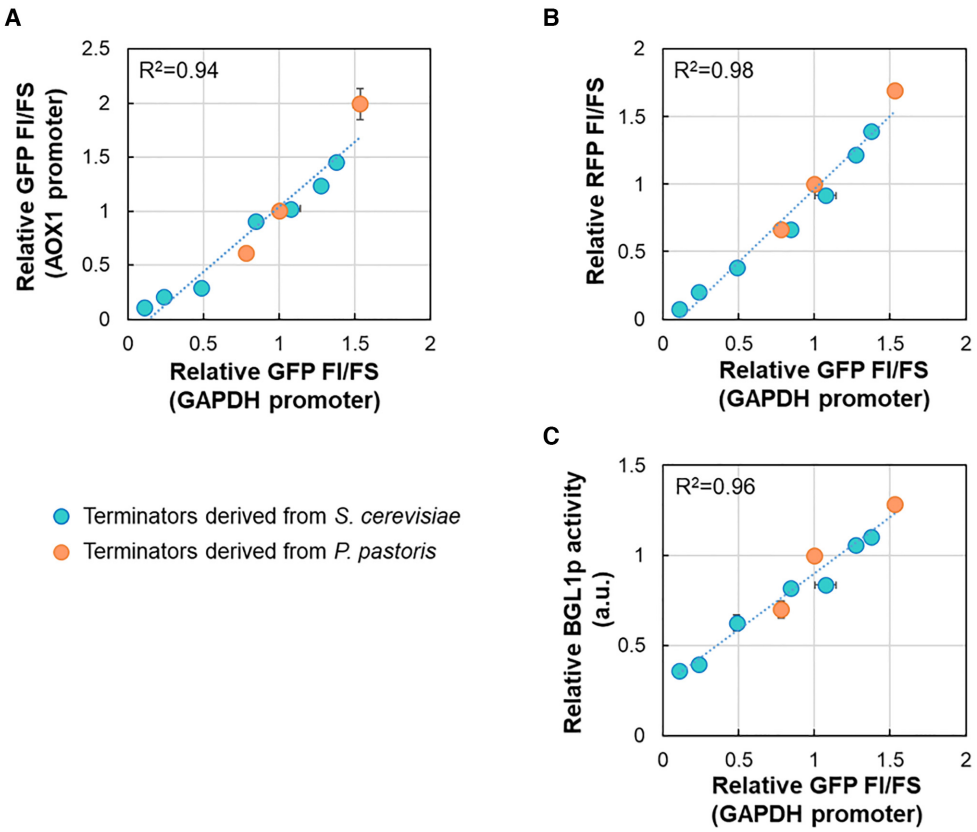


Figure 5. Effect of promoter and reporter gene exchange on terminator activities in *Pichia pastoris*. Seven terminators derived from *Saccharomyces cerevisiae* (blue) and three terminators from *P. pastoris* (orange) were used. Terminator activities (relative GFP FI/FS) of all strains are relative to that of the *Pp_GAPDH* terminator strain. Terminator activities of promoter and reporter gene-exchanged strains were compared with that of *GAPDH* promoter-EGFP-terminator strains. (**A**) Effect of promoter exchange on terminator activity. The horizontal and vertical axes show the terminator activity for the *GAPDH* and *AOX1* promoter strains, respectively. (**B** and **C**) Effects on terminator activity of reporter gene exchange to RFP (E2Crimson) (B) or *Rhizopus oryzae* glucoamylase secretion signal-BGL1p (C). The horizontal axis shows terminator activity (relative GFP FI/FS). The vertical axis shows normalized RFP FI (relative RFP FI/FS) (B) and relative BGL1p activity (C). Error bars represent the standard deviation from at least three biological replicates.

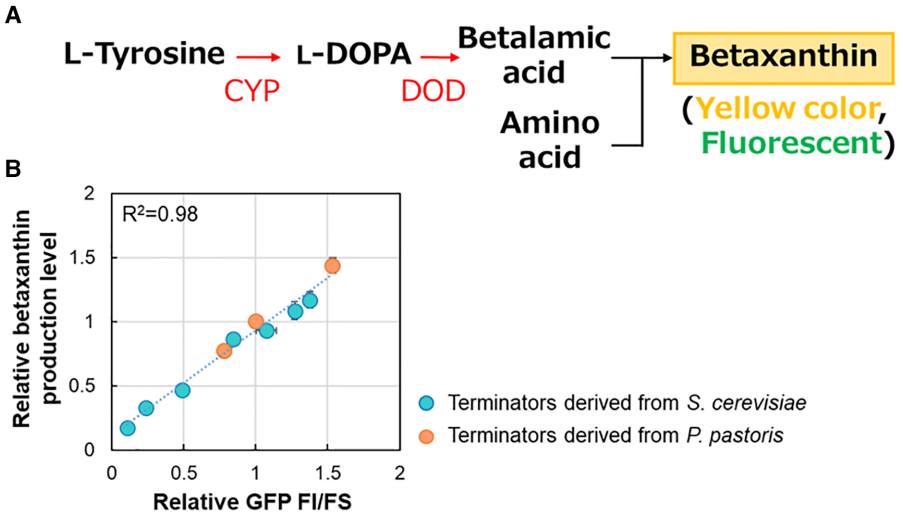


Figure 6. Tuning of betaxanthin production using a terminator exchange strategy. (**A**) The betaxanthin production pathway in yeast. A tyrosine hydroxylase (CYP76AD1) with a double mutation (W13L/F309L) converts L-tyrosine into L-DOPA, and a DOPA deoxygenase (DOD) converts L-DOPA into betalamic acid. The betalamic acid reacts with amino acids in yeast cells to form the fluorescent yellow pigment betaxanthin. (**B**) Effect of terminator exchange at the CYP76AD1 gene. Ten terminators were used, as indicated in Figure 5. Horizontal axis, terminator activity (relative GFP FI/FS); vertical axis, FI of the culture supernatant (relative betaxanthin levels). Betaxanthin production levels of all strains are shown relative to levels observed for the *Pp_GAPDH* terminator strain. Error bars represent the standard deviation from at least three biological replicates.

data indicate that DOD expression was at levels sufficient to produce betaxanthin, at even at a low copy number of DOD transgenes. Hereafter, a strain carrying two copies of the DOD transgene was used for further study.

To investigate whether metabolic pathway flux can be tuned using the terminator-exchange strategy, we next constructed betaxanthin-producing strains carrying the DOD transgene by introducing versions of the CYP76AD1 transgene with the 10 yeast terminators as in the swap experiments, then evaluated betaxanthin levels. As shown in Figure 6B and Supplementary Table S5, betaxanthin levels were different among the terminator strains: the highest was observed for the *Pp_AOX1*t strain and the lowest for the *Sc_ICY2*t strain, with an overall range of about 10-fold. Consistent with the swap experiments described above, we found that betaxanthin production by each terminator strain also correlates with terminator activity as measured by cytoplasmic GFP levels in *P. pastoris* (Figure 2). As we had predicted, these results demonstrate that we can tune metabolic pathway flux just by changing the terminator sequence in *P. pastoris*.

Terminators influence mRNA abundance and mRNA half-life in *P. pastoris*

As mentioned above, terminator activities in *P. pastoris* were not dependent on promoter or reporter ORF sequences, but were dependent on nucleotide sequence. Where does the difference in terminator activity come from? To address this, we next investigated the mechanism(s) underlying differences in terminator activity using the GFP expression construct under the control of the *GAPDH* promoter and the 10 yeast terminators used in the swap and metabolic pathway flux experiments.

We first checked the growth rate of the terminator strains; however, we found that the growth rates of the 10 terminator strains were the same in BMGY (glycerol) culture conditions (Supplementary Figure S3). Subsequently, to evaluate mRNA abundance, the GFP mRNA level in each terminator strain was measured using reverse transcription quantitative PCR (RT-qPCR) analysis. As shown in Figure 7 and Supplementary Figure S4A, we found that except for *Pp_AOX1*t, GFP mRNA levels were different for different terminator strains and correlated with terminator activity. This is the same trend shown for *S. cerevisiae*, in

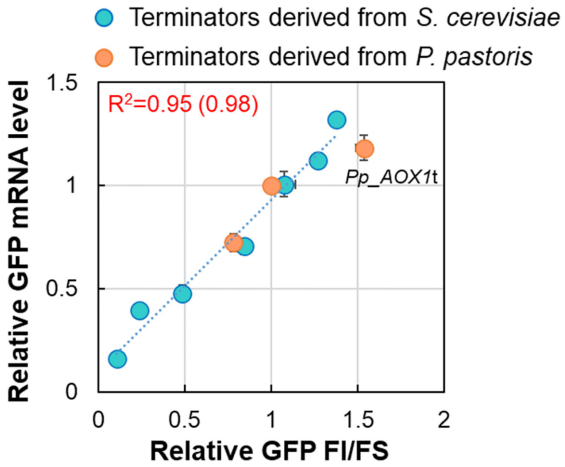


Figure 7. EGFP mRNA levels of *P. pastoris* strains expressing EGFP under the control of the *GAPDH* promoter with various terminators. Ten terminators described in Figures 5 and 6 were used in this analysis. The EGFP mRNA levels of all strains are shown relative to levels observed for the *Pp_GAPDH* terminator strain. Error bars represent standard deviation from three biological replicates. Terminator activity (relative GFP FI/FS) values are the same as in Figure 2. The R-squared value and that in parentheses were calculated from all 10 terminators or from the nine terminators except for *Pp_AOX1*t, respectively.

which mRNA levels for a fluorescent protein (yECitrine) expressed by the terminator test strains with different terminator activities correlated with FI (14).

Finally, to examine whether the mRNA level reflected the stability (degradability) of the mRNA, the mRNA half-life of each GFP mRNA in the different terminator strains was measured using a method described previously (36) with some modifications. Namely, qPCR was used instead of northern blot analysis to quantify the remaining GFP mRNA level in cells (36). To evaluate mRNA half-life, we used the EGFP-terminator strains in which transcription is driven by the methanol inducible *AOX1* promoter; we note that GFP fluorescence levels in these strains correlate with levels observed using the *GAPDH* promoter (Figure 5A). After the strains were cultured with BMMY (methanol) medium to achieve steady expression of the *EGFP* mRNA, transcription was turned off by switching the media from methanol to glucose, and we measured the mRNA degradation over time by qPCR.

For six of the ten terminator strains, we successfully determined the mRNA half-life. For the other four terminator strains, which had weaker activities, we could not determine half-life because their half-lives were too short to be measured accurately (Figure 8). Specifically, even the highest terminator strain (*Pp_AOX1*t) exhibited an mRNA half-life of only 7.3 min, and the weakest terminator strain (*Sc_ICY2*t) never displayed detectable mRNA even at the $t = 0$ sampling point. Surprisingly, the mRNA half-lives in *P. pastoris* were much shorter than those in *S. cerevisiae;* the *Sc_CYC1*t in *P. pastoris* (about 2 min) indicated about five-times shorter than that in *S. cerevisiae* (about 10 min) (14) and the maximum mRNA half-life in *S. cerevisiae* reported was about 27 min (*PRM9*t) (14). In this small range of half-life time, the mRNA half-life of the six terminators measured correlated well with GFP fluorescence ($R^2 = 0.98$, Figure 8A). Although the correlation between the mRNA half-life and the GFP mRNA level was small ($R^2 = 0.64$), the use of five terminators, excluding the *Pp_AOX1*t for analysis, showed a higher correlation ($R^2 = 0.87$) than when all six terminators were used (Figure 8B). Together, these indicated that, with the exception of *Pp_AOX1*t, which may be regulated differently, terminator activity was influenced by mRNA abundance and half-life in *S. cerevisiae*, consistent with the results of a previous study (14).

Terminator activity in *P. pastoris* correlates with stability of 3′-UTR secondary structure

Having observed that the terminator activity correlated with mRNA abundance and half-life, we explored the possible causes of differences in the mRNA abundance and half-life. Specifically, we investigated whether these differences in *P. pastoris* resulted from differences in the stability of RNA secondary structure in the 3′-UTR, as stability in the 3′-UTR could reduce mRNA degradation by blocking access by nucleases. Past studies have used MFE as an indicator of RNA secondary structure stability (41,42). A lower MFE indicates the presence of a more stable RNA secondary structure. Importantly, when we compare the MFE of different sequences, the length of the sequences should be the same, as MFE strongly depends on the length of a nucleotide sequence; e.g. a longer sequence generally has a lower MFE (43). In this study, we therefore extracted fixed length sequences downstream of the stop codons and compared MFE for these sequences.

To confirm the range of 3′-UTR length to be searched in the 3′-UTR stability analysis, we firstly identified the 3′-UTR sequences and length for each of the 10 terminators by 3′-RACE (Supplementary Table S6). The averaged lengths of the 3′-UTRs ranged from 57 to 173 nucleotides including the XhoI recognition site positioned next to the stop codon of the upstream GFP-encoding ORF (Figure 1). Based on this analysis, we found that the length of the 3′-UTR does not correlate with GFP mRNA level or GFP FI suggesting that 3′-UTR length is unrelated to the effect of a terminator on expression (Supplementary Figure S5).

Next, we calculated the MFE of the fixed length 3′-UTR sequence using the RNAfold program (37), then compared these values with the mRNA levels of the 10 terminators. GC content was also calculated for each 3′-UTR sequence. In this analysis, only the 3′-UTR sequence from 20 to 100 nt was considered, as the 3′-UTR length for the 10 terminators used in this analysis was typically around 100 nt. We observed a strong inverse correlation between MFE and mRNA level when the length of the 3′-UTR was above 50 nt (Figure 9A). A maximum inverse correlation was observed when the UTR length was 80–84 nt (Spearman's rho $= -0.976$) (Figure 9B). To evaluate the statistical significance of this inverse correlation, we calculated the *P*-value of the minimum rho value based on the empirical distribution obtained from randomly shuffled data (see 'Materials and Methods' section). We observed a *P*-value of 0.0014, indicating that the inverse correlation is statistically significant. By contrast, we observed that the GC content and nucleosome occupancy score showed weaker correlations with GFP mRNA levels at every 3′-UTR length (Figure 9A), suggesting that terminator activity is affected by the stability of 3′-UTR secondary structure but not GC content and nucleosome occupancy. We also conducted the same analysis using mRNA half-life instead of mRNA level, and observed the same trend (Supplementary Figure S6).

In a test of all terminators in the catalog except the synthetic terminators, we next asked if there is a correlation between terminator activity (GFP expression level) and 3′-UTR secondary structure stability. As shown in Figure 9C, the MFE and terminator activities were inversely correlated, consistent with the data shown in Figure 9A. The strongest inverse correlation was observed when the length of the 3′-UTR was 70 nt (Spearman's rho $= -0.48$, *P* value $= 3.6 \times 10^{-4}$) (Figure 9D). GC content and nucleosome occupancy showed weaker inverse correlations (Figure 9C).

DISCUSSION

In this study, we successfully characterized 72 terminators, including terminator sequences derived from *S. cerevisiae* or *P. pastoris*, and synthetic sequences, resulting in a terminator catalog for *P. pastoris*. The degree of terminator-based tunability observed in *P. pastoris* was about 17-fold. We also demonstrated that the terminators tend to work independently on upstream sequences, including promoter and ORF sequences, as has been observed for *S. cerevisiae*

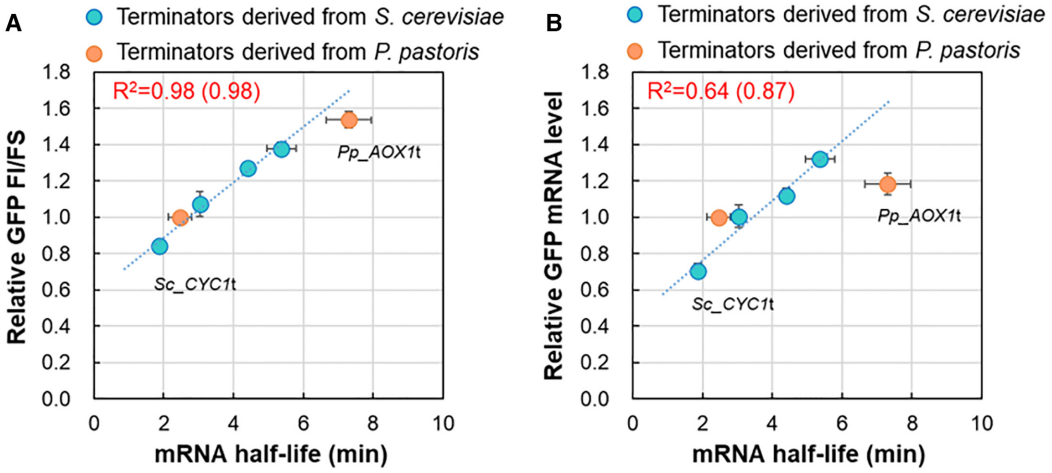


Figure 8. mRNA half-life analysis of *Pichia pastoris* strains. Yeast strains expressing EGFP under the control of the *AOX1* promoter were used for the analysis. The mRNA half-life was successfully determined for six out of the ten terminators in *P. pastoris*. (**A**) mRNA half-life time versus terminator activity (relative GFP FI/FS). Terminator activity values are the same as in Figure 2. (**B**) mRNA half-life time versus EGFP mRNA level. The EGFP mRNA levels are the same as in Figure 7. Error bars in the mRNA half-life analysis represent standard deviation for three technical replicates. The R-squared value and that in parentheses were calculated from all six terminators and the five terminators except for *Pp_AOX1*t, respectively.

(13). Thus, we can tune the expression of target proteins in *P. pastoris* within a range of more than 10-fold simply by choosing an appropriate terminator sequence. As part of this study, we characterized for the first time the activity in *P. pastoris* of synthetic terminators originally developed for use in *S. cerevisiae*. Although the synthetic terminators resulted in high expression in *S. cerevisiae* (18,19), we did not observe this effect for *P. pastoris*. This suggests that different mechanisms of regulation for protein expression are at play in *S. cerevisiae* versus *P. pastoris*. Whereas the promoter-exchange strategies have been extensively tested, only a few past studies explored use of a terminator-exchange strategy to tune protein expression levels (15). In *S. cerevisiae*, the activity of a given promoter is often dependent on culture conditions (11,12). By contrast, it was reported that terminators work independently of culture conditions, in a study that explored five different carbon sources (16). Thus, terminators should be useful tools for tuning protein expression levels in a manner that is robust and tractable under any culture condition. Therefore, the terminator catalog created here will be useful for applications in *P. pastoris* that require a need to control the protein expression levels within a certain range.

In our analysis of 72 terminators, we showed that *S. cerevisiae* terminator sequences had a high degree of transferability to *P. pastoris*. This has a merit for correct integration into a target locus on the *P. pastoris* genome. It was previously reported that there is a large clonal variability between *P. pastoris* transformants due to ectopic integration events within the host genome (44). To avoid ectopic integration, sequence similarity of the transgene to the host genome should be decreased (45). Most yeast promoters do not work in heterologous environments (i.e. do not work in other yeast strains). By contrast, use of a heterologous terminator provides a means for reducing the chance for undesired integrations of the transgene into the host genome. Thus, the high degree of transferability of the *S. cerevisiae* terminator sequences is valuable for integration on *P. pastoris*; moreover, this approach might also be applied in other yeast species and strains.

In this study, we observed that terminator activity (GFP expression level) correlated with mRNA level and half-life (Figures 7 and 8). This motivated us to explore the possibility that stable RNA secondary structures in the 3′-UTR prevent the mRNA from being degraded, and hence increase mRNA expression levels and mRNA half-lives. We observed a statistically significant correlation between the MFEs of 3′-UTRs and terminator activity (Figure 9). When GC content and the nucleosome occupancy score were used to calculate correlation instead of MFE, the correlation was reduced, suggesting that differences in RNA secondary structure rather than in GC content and nucleosome occupancy are the likely source of differences in terminator activity. Although further work needs to be done to fully demonstrate a contribution by RNA secondary structure to terminator activity, our analysis provides preliminary evidence in the form of a statistically significant correlation between MFE and terminator activity in *P. pastoris*.

Our results suggest that *Pp_AOX1*t is regulated via a different mechanism as compared with the other the terminators characterized in this study. We reason that there is a possibility that *Pp_AOX1*t might be regulated by endogenous *trans*-acting factors, additional to Nab6p and Pap1p that act on the *DIT1* terminator in *S. cerevisiae* (17). Firstly, terminator activity of *Pp_AOX1*t increased over culture time (Figure 3), which is the same trend observed for the endogenous *DIT1* terminator in *S. cerevisiae* (16). Secondly, *Pp_AOX1*t was an outlier with regards to a correlation between the GFP mRNA level and terminator activity or mRNA half-life (Figures 7 and 8; Supplementary Table S4). Indeed, although GFP mRNA level in the *Pp_AOX1*t strain was lower than that of the *Sc_RPL3*t strain (Supplementary Figure S4A), terminator activity of the endogenous *Pp_AOX1*t appeared to be higher than that of the het-

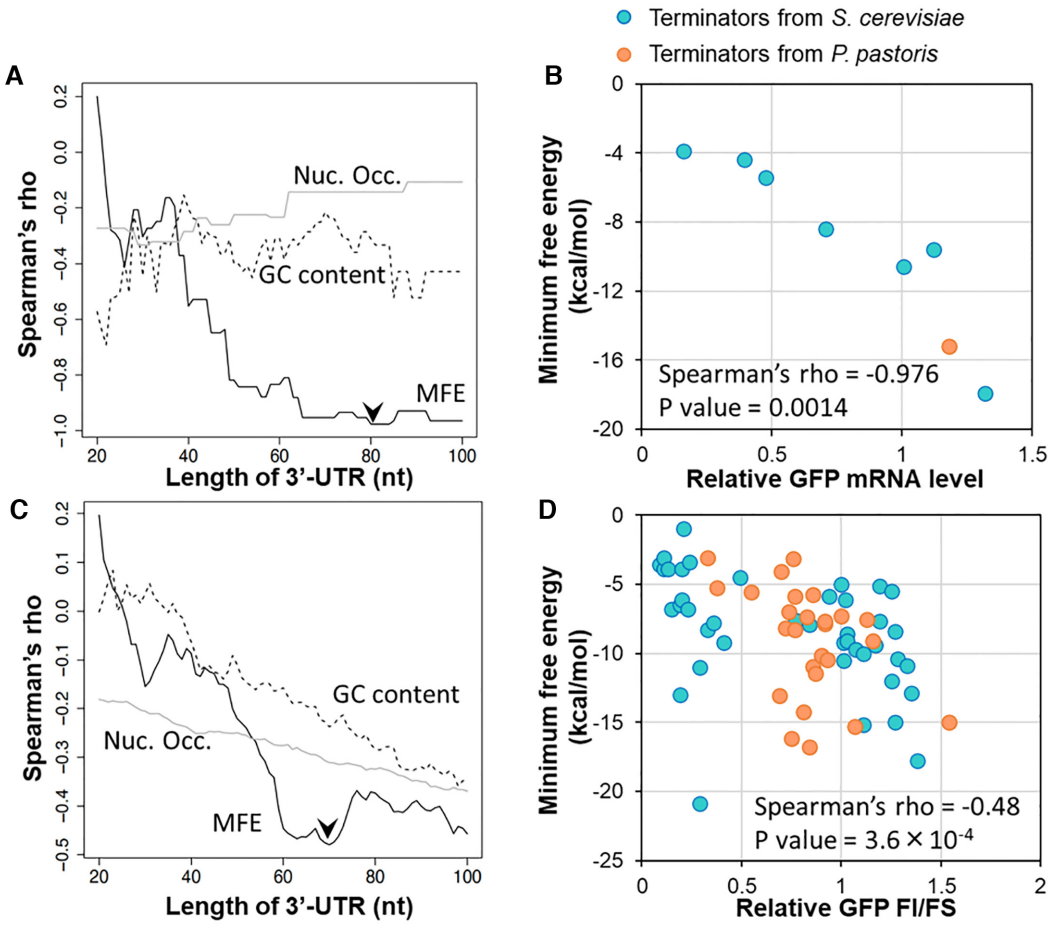


Figure 9. *In silico* analysis of stability of 3′-UTR secondary structure and terminator activity. (**A**) Spearman's rho between the MFE and mRNA level for the 10 terminators (solid line). *X*-axis, the length of 3′-UTR used for the calculation of MFE. An arrowhead indicates the length with the lowest rho value (80 nt). Dashed line and gray line, the rho values when GC content and nucleosome occupancy score are used instead of MFE, respectively. (**B**) A plot of MFEs and mRNA levels at a UTR length of 80 nt. Two of the ten terminators were not included because their 3′-UTRs are <80 nt. (**C**) The same analysis as in (A) when all 69 terminators are included. In this panel, terminator activity values (relative GFP FI/FS) were used instead of mRNA levels. An arrowhead indicates the length with the lowest rho value (70 nt). Note that the range shown on the *Y*-axis is different from that of (A). (**D**) A plot of MFE and terminator activity values at a UTR length of 70 nt. Two of 69 terminators were not included for the same reason as in (B).

erologous *Sc_RPL3*t in *P. pastoris* (Figure 2 and Supplementary Table S4). However, the half-life of mRNA in the *Pp_AOX1*t strain was higher than that in the *Sc_RPL3*t strain (Supplementary Figure S4B). Thus, *trans*-acting factors might be engaged in extension in the *Pp_AOX1*t strain of the mRNA half-life, which was longer than expected from the mRNA abundance (Figure 8B), although further analysis is needed.

In this study, we demonstrated that using a terminator-exchange strategy to control protein expression levels is effective in *P. pastoris*, consistent with what had previously been shown for *S. cerevisiae*. In this first analysis, in which we tested 72 terminators, we were able to achieve a tunable range of expression in *P. pastoris* of 17-fold. To further expand the range of the terminator activities in *P. pastoris*, terminator sequences with much higher and lower activities could be identified in future work based on endogenous or heterogeneous yeast genomes, or through synthetic design. Altogether, combining the terminator-exchange strategy with a promoter-exchange strategy will expand the range and precision of metabolic engineering and synthetic biology applications that can be performed using *P. pastoris*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Yuri Sakihama for helpful discussions.

FUNDING

Japan Agency for Medical Research and Development (AMED) under the project focused on Developing Key Technology for Discovering and Manufacturing Drugs for Next-Generation Treatment and Diagnosis [JP19ae0101055, JP19ae0101060]; New Energy and Industrial Technology Development Organization

(NEDO) under the project for Development of Production Techniques for Highly Functional Biomaterials Using Smart Cells of Plants and Other Organisms (Smart Cell Project) [P16009]; Japan Science and Technology Agency (JST) under the JST-Mirai Program [JPMJMI17EJ]. Funding for open access charge: AMED.
Conflict of interest statement. None declared.

REFERENCES

1. Nielsen,J. and Keasling,J.D. (2016) Engineering cellular metabolism. *Cell*, **164**, 1185–1197.
2. Lian,J., Hamedirad,M., Hu,S. and Zhao,H. (2017) Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.*, **8**, 1688.
3. Besada-Lombana,P.B., McTaggart,T.L. and Da Silva,N.A. (2018) Molecular tools for pathway engineering in Saccharomyces cerevisiae. *Curr. Opin. Biotechnol.*, **53**, 39–49.
4. Petersen,S.D., Zhang,J., Lee,J.S., Jakociunas,T., Grav,L.M., Kildegaard,H.F., Keasling,J.D. and Jensen,M.K. (2018) Modular 5′-UTR hexamers for context-independent tuning of protein expression in eukaryotes. *Nucleic Acids Res.*, **46**, e127.
5. Hoshida,H., Kondo,M., Kobayashi,T., Yarimizu,T. and Akada,R. (2017) 5′-UTR introns enhance protein expression in the yeast Saccharomyces cerevisiae. *Appl. Microbiol. Biotechnol.*, **101**, 241–251.
6. Alper,H., Fischer,C., Nevoigt,E. and Stephanopoulos,G. (2005) Tuning genetic control through promoter engineering,. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 12678–12683.
7. Lee,M.E., Aswani,A., Han,A.S., Tomlin,C.J. and Dueber,J.E. (2013) Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.*, **41**, 10668–10678.
8. Hartner,F.S., Ruth,C., Langenegger,D., Johnson,S.N., Hyka,P., Lin-Cereghino,G.P., Lin-Cereghino,J., Kovar,K., Cregg,J.M. and Glieder,A. (2008) Promoter library designed for fine-tuned gene expression in Pichia pastoris. *Nucleic Acids Res.*, **36**, 1–15.
9. Stadlmayr,G., Mecklenbräuker,A., Rothmüller,M., Maurer,M., Sauer,M., Mattanovich,D. and Gasser,B. (2010) Identification and characterisation of novel Pichia pastoris promoters for heterologous protein production. *J. Biotechnol.*, **150**, 519–529.
10. Vogl,T., Sturmberger,L., Kickenweiz,T., Wasmayer,R., Schmid,C., Hatzl,A.M., Gerstmann,M.A., Pitzer,J., Wagner,M., Thallinger,G.G. *et al.* (2016) A toolbox of diverse promoters related to methanol utilization: functionally verified parts for heterologous pathway expression in pichia pastoris. *ACS Synth. Biol.*, **5**, 172–186.
11. Peng,B., Williams,T.C., Henry,M., Nielsen,L.K. and Vickers,C.E. (2015) Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microb. Cell Fact.*, **14**, 91.
12. Yuan,T., Guo,Y., Dong,J., Li,T., Zhou,T., Sun,K., Zhang,M., Wu,Q., Xie,Z., Cai,Y. *et al.* (2017) Construction, characterization and application of a genome-wide promoter library in Saccharomyces cerevisiae. *Front. Chem. Sci. Eng.*, **11**, 107–116.
13. Yamanishi,M., Ito,Y., Kintaka,R., Imamura,C., Katahira,S., Ikeuchi,A., Moriya,H. and Matsuyama,T. (2013) A genome-wide activity assessment of terminator regions in saccharomyces cerevisiae provides a 'terminatome' toolbox. *ACS Synth. Biol.*, **2**, 337–347.
14. Curran,K.A., Karim,A.S., Gupta,A. and Alper,H.S. (2013) Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab. Eng.*, **19**, 88–97.
15. Matsuyama,T. (2019) Recent developments in terminator technology in Saccharomyces cerevisiae. *J. Biosci. Bioeng.*, **128**, 655–661.
16. Ito,Y., Yamanishi,M., Ikeuchi,A., Imamura,C., Tokuhiro,K., Kitagawa,T. and Matsuyama,T. (2013) Characterization of five terminator regions that increase the protein yield of a transgene in Saccharomyces cerevisiae. *J. Biotechnol.*, **168**, 486–492.
17. Ito,Y., Kitagawa,T., Yamanishi,M., Katahira,S., Izawa,S., Irie,K., Furutani-Seiki,M. and Matsuyama,T. (2016) Enhancement of protein production via the strong DIT1 terminator and two RNA-binding proteins in Saccharomyces cerevisiae. *Sci. Rep.*, **6**, 36997.
18. Curran,K.A., Morse,N.J., Markham,K.A., Wagman,A.M., Gupta,A. and Alper,H.S. (2015) Short synthetic terminators for improved heterologous gene expression in yeast. *ACS Synth. Biol.*, **4**, 824–832.
19. Ahmed,M.S., Ikram,S., Rasool,A. and Li,C. (2019) Design and construction of short synthetic terminators for B-amyrin production in Saccharomyces cerevisiae. *Biochem. Eng. J.*, **146**, 105–116.
20. Gasser,B., Prielhofer,R., Marx,H., Maurer,M., Nocon,J., Steiger,M., Puxbaum,V., Sauer,M. and Mattanovich,D. (2013) Pichia pastoris: protein production host and model organism for biomedical research. *Futur. Microbiol.*, **8**, 191–208.
21. Ahmad,M., Hirz,M., Pichler,H. and Schwab,H. (2014) Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.*, **98**, 5301–5317.
22. Peña,D.A., Gasser,B., Zanghellini,J., Steiger,M.G. and Mattanovich,D. (2018) Metabolic engineering of Pichia pastoris. *Metab. Eng.*, **50**, 2–15.
23. Schwarzhans,J.P., Luttermann,T., Geier,M., Kalinowski,J. and Friehs,K. (2017) Towards systems metabolic engineering in Pichia pastoris. *Biotechnol. Adv.*, **35**, 681–710.
24. Wagner,J.M. and Alper,H.S. (2016) Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. *Fungal Genet. Biol.*, **89**, 126–136.
25. Pesole,G., Sabino,L., Grillo,G., Licciulli,F., Mignone,F., Gissi,C. and Saccone,C. (2002) UTRdb and UTRsite: specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs. Update 2002. *Nucleic Acids Res.*, **30**, 335–340.
26. Jacobs,G.H., Chen,A., Stevens,S.G., Stockwell,P.A., Black,M.A., Tate,W.P. and Brown,C.M. (2009) Transterm: a database to aid the analysis of regulatory sequences in mRNAs. *Nucleic Acids Res.*, **37**, 72–76.
27. Barreau,C., Paillard,L. and Osborne,H.B. (2005) AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.*, **33**, 7138–7150.
28. Tanguay,R.L. and Gallie,D.R. (1996) Translational efficiency is regulated by the length of the 3′ untranslated region. *Mol. Cell. Biol.*, **16**, 146–156.
29. Imamachi,N., Salam,K.A., Suzuki,Y. and Akimitsu,N. (2017) A GC-rich sequence feature in the 3′ UTR directs UPF1-dependent mRNA decay in mammalian cells. *Genome Res.*, **27**, 407–418.
30. Wu,X. and Bartel,D.P. (2017) Widespread influence of 3′-End structures on mammalian mRNA processing and stability. *HHS Public Access*, **169**, 905–917.
31. Sambrook,J. and Russel,D. (2001) In: *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Press, NY.
32. Ito,Y., Watanabe,T., Aikawa,S., Nishi,T., Nishiyama,T., Nakamura,Y., Hasunuma,T., Okubo,Y., Ishii,J. and Kondo,A. (2018) Deletion of DNA ligase IV homolog confers higher gene targeting efficiency on homologous recombination in Komagataella phaffii. *FEMS Yeast Res.*, **18**, doi:10.1093/femsyr/foy074.
33. Nakamura,Y., Nishi,T., Noguchi,R., Ito,Y., Watanabe,T., Nishiyama,T., Aikawa,S., Hasunuma,T., Ishii,J., Okubo,Y. *et al.* (2018) A stable, autonomously replicating plasmid vector containing Pichia pastoris centromeric DNA. *Appl. Environ. Microbiol.*, **84**, e02882-17.
34. Strack,R.L., Hein,B., Bhattacharyya,D., Hell,S.W., Keenan,R.J. and Glick,B.S. (2009) A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry*, **48**, 8279–8281.
35. Inokuma,K., Bamba,T., Ishii,J., Ito,Y., Hasunuma,T. and Kondo,A. (2016) Enhanced cell-surface display and secretory production of cellulolytic enzymes with Saccharomyces cerevisiae Sed1 signal peptide. *Biotechnol. Bioeng.*, **113**, 2358–2366.
36. Coller,J. (2008) In: *Methods to Determine mRNA Half-Life in Saccharomyces cerevisiae*. 1st edn. Elsevier Inc, Amsterdam.
37. Lorenz,R., Bernhart,S.H., Siederdissen,C.H. zu, Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorith. Mol. Biol.*, **6**, 26.
38. Morse,N.J., Gopal,M.R., Wagner,J.M. and Alper,H.S. (2017) Yeast terminator function can be modulated and designed on the basis of predictions of nucleosome occupancy. *ACS Synth. Biol.*, **6**, 2086–2095.
39. Xi,L., Fondufe-Mittendorf,Y., Xia,L., Flatow,J., Widom,J. and Wang,J.P. (2010) Predicting nucleosome positioning using a duration Hidden Markov Model. *BMC Bioinformatics*, **11**, 346.

40. Deloache,W.C., Russ,Z.N., Narcross,L., Gonzales,A.M., Martin,V.J.J. and Dueber,J.E. (2015) An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat. Chem. Biol.*, **11**, 465–471.
41. Cambray,G., Guimaraes,J.C. and Arkin,A.P. (2018) Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in Escherichia coli. *Nat. Biotechnol.*, **36**, 1005–1015.
42. Loong,S.N.K. and Mishra,S.K. (2007) Unique folding of precursor microRNAs: quantitative evidence and implications for de novo identification. *Bioinformatics*, **13**, 170–187.
43. Trotta,E. (2014) On the normalization of the minimum free energy of RNAs by sequence length. *PLoS One*, **9**, e113380.
44. Schwarzhans,J.-P., Wibberg,D., Winkler,A., Luttermann,T., Kalinowski,J. and Friehs,K. (2016) Non-canonical integration events in Pichia pastoris encountered during standard transformation analysed with genome sequencing. *Sci. Rep.*, **6**, 38952.
45. Schwarzhans,J.P., Wibberg,D., Winkler,A., Luttermann,T., Kalinowski,J. and Friehs,K. (2016) Integration event induced changes in recombinant protein productivity in Pichia pastoris discovered by whole genome sequencing and derived vector optimization. *Microb. Cell Fact.*, **15**, 84.