One-Step In Vivo Assembly of Multiple DNA Fragments and Genomic Integration in Komagataella phaffii

Nishi, Teruyuki ; Ito, Yoichiro ; Nakamura, Yasuyuki ; Yamaji, Taiki ; Hashiba, Noriko ; Tamai, Masaya ; Yasohara, Yoshihiko ; Ishii, Jun ;…

(Citation)
ACS Synthetic Biology,11(2):644-654

(Issue Date)
2022-01-30

(Resource Type)
journal article

(Version)
Version of Record

(URL)
https://hdl.handle.net/20.500.14094/0100481886

ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

One-Step *In Vivo* Assembly of Multiple DNA Fragments and Genomic Integration in *Komagataella phaffii*

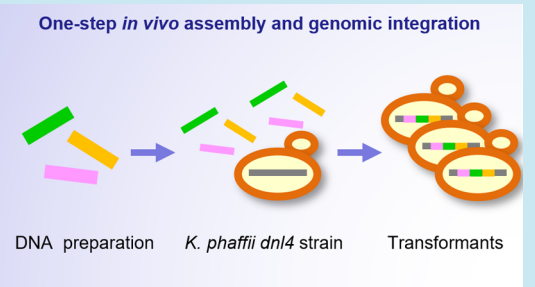
Teruyuki Nishi, Yoichiro Ito, Yasuyuki Nakamura, Taiki Yamaji, Noriko Hashiba, Masaya Tamai, Yoshihiko Yasohara, Jun Ishii,* and Akihiko Kondo

Cite This: *ACS Synth. Biol.* 2022, 11, 644–654

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The methylotrophic yeast species *Komagataella phaffii* (synonym: *Pichia pastoris*) is widely used as a host for recombinant protein production. Although several genetic engineering techniques are being employed on *K. phaffii*, advanced methods such as *in vivo* DNA assembly in this yeast species are required for synthetic biology applications. In this study, we established a technique for accomplishing one-step *in vivo* assembly of multiple DNA fragments and genomic integration in *K. phaffii*. To concurrently achieve an accurate multiple DNA assembly and a high-efficient integration into the target genomic locus *in vivo*, a *K. phaffii* strain, lacking a non-homologous end joining-related protein, DNA ligase IV (Dnl4p), that has been reported to improve gene targeting efficiency by homologous recombination, was used. Using green fluorescent protein along with the lycopene biosynthesis, we showed that our method that included a Dnl4p-defective strain permits direct and easy engineering of *K. phaffii* strains.



KEYWORDS: *Komagataella phaffii, in vivo DNA assembly, genomic integration, homologous recombination (HR), non-homologous end joining (NHEJ)*

INTRODUCTION

The methylotrophic yeast *Komagataella phaffii*, also known under the synonym *Pichia pastoris*, has been widely used to produce valuable recombinant proteins, including enzymes and biologics for academic and industry use.[1−4] *K. phaffii* has been recently reported to be also used as a cell factory to produce a variety of chemicals.[5] Particularly, this species has an advantage of being convenient for large-scale high-density culturing as well as secretory protein production that is achieved by employing strong promoters derived from genes with either methanol-inducible or constitutive expression such as alcohol oxidase 1 (*AOX1*) and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). Various genetic tools have been developed for *K. phaffii*,[6−8] but not as much as for the budding yeast *Saccharomyces cerevisiae*. For example, a multiloci integration of *in vivo* assembled DNA fragments as represented by the CasEMBLR system[9] was reported in *S. cerevisiae*; however, it is still challenging in *K. phaffii*. To promote applications of this yeast into the bioeconomy, it is necessary to develop more advanced genetic engineering and synthetic biology tools for *K. phaffii* such as the CasEMBLR, which has a potential to make applications for combinatorial optimization of gene expressions and genetic resources.

Detailed genome information of *K. phaffii*, obtained by next-generation sequencing,[10−12] has accelerated the evolvement of genetic engineering technologies, *e.g.*, several autonomously replicating plasmid vectors have been developed using either autonomously replicating sequences or centromeric DNA from this species.[13−17] Although autonomously replicating plasmids have been shown to be a favorable tool in transient expression verification, vector systems allowing genomic integration represent an essential and a mainstream approach to stably produce recombinant proteins in a long-term fashion.[18−20] However, a conventional method for genomic integration[21,22] usually needs several steps (*i.e.*, *in vitro* circular plasmid construction by either ligation or DNA assembly, transformation of *Escherichia coli* and multiplication of the target DNA, subsequent linearization of the circular DNA, and finally, transformation to *K. phaffii*) and, therefore, it is still both cumbersome and time-consuming to obtain stable expression strains of *K. phaffii* (Figure 1).

For easier construction and stable development of *K. phaffii* strains, *in vivo* DNA assembly systems of *S. cerevisiae* that are based on a gap-repair cloning procedure, so-called "yeast assembly",[23−26] would represent a good model system. The yeast assembly is based on the homologous recombination (HR) and often utilized in *S. cerevisiae* for connecting

Received: July 12, 2021
Published: January 31, 2022

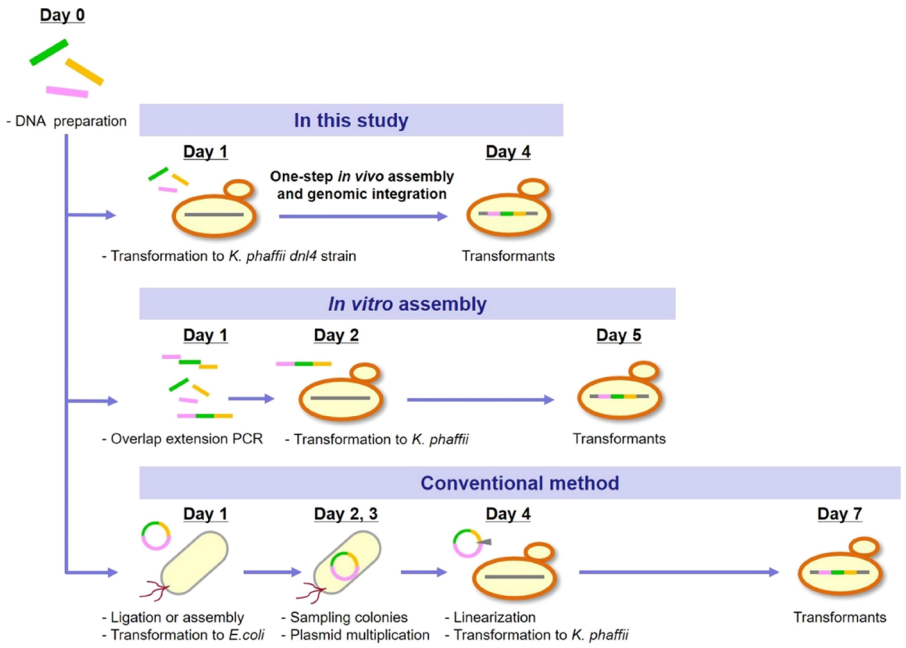


Figure 1. Schematic representation of the one-step *in vivo* assembly of multiple DNA fragments and genomic integration in *K. phaffii*. For comparison, the method that uses *in vitro* assembly of multiple DNA fragments and the conventional method are also illustrated. These methods need multiple steps: overlap extension PCR (OE-PCR), *in vitro* circular plasmid construction by either ligation or DNA assembly, plasmid multiplication via *E. coli* transformation, linearization of the circular DNA, and transformation to *K. phaffii*. On the other hand, the method presented in this study allows one-step construction to obtain *K. phaffii* transformants correctly integrating the assembled DNA fragments at the target locus.

(assembling) recombinant DNA fragments with short homologous sequences as mutual fragments. To date, the overlap extension PCR (OE-PCR) technique has been utilized *in vitro* for cloning multiple DNA fragments into *K. phaffii* (Figure 1).[27,28] Several methods applicable for *in vivo* DNA assembly in *K. phaffii* have been also reported; however, they were developed for inserting DNA fragments into autonomously replicating vectors.[14,29,30] To the best of our knowledge, there are no efficient genetic tools reported for enabling direct genomic integration concurrently with *in vivo* assembly of multiple DNA fragments in *K. phaffii*.

Since *S. cerevisiae* has HR efficiency higher than nonhomologous end joining (NHEJ), it innately possesses a potential for both *in vivo* assembly of multiple DNA fragments and genomic integration.[31] In contrast, *K. phaffii* tends to randomly integrate a linear DNA fragment into unexpected genome loci (even when a single (not multiple) DNA fragment is integrated) due to the dominance of NHEJ.[32] For precise genomic integration into a target locus, NHEJ-related proteins such as the telomeric Ku complex subunit Ku70p have been often deleted through genetic engineering.[33] Recently, we reported that deletion of another NHEJ-related protein, DNA ligase IV (Dnl4p), further improved HR efficiency of genomic integration into the target loci of *K. phaffii*.[34] More recently, Cai *et al.* reported that overexpressing *RAD52* and deleting *MPH1* enhanced HR in combination with the CRISPR-Cas9 system.[35] The development of these protocols has an important implication for potentiating *in vivo* genomic integration and assembly of multiple DNA fragments.

In this study, we intended to develop a genetic engineering platform for *K. phaffii* that enables both one-step *in vivo* assembly of multiple DNA fragments and genomic integration. To establish this methodology, we utilized the NHEJ-defective *K. phaffii dnl4* strain having a higher HR activity than the wild-type strain. Using green fluorescent protein (GFP) along with lycopene biosynthesis, we aimed to show if the Dnl4p-defective strain permits efficient and concurrent *in vivo* DNA assembly and genomic integration that would accelerate genetic engineering of *K. phaffii* strains.

RESULTS AND DISCUSSION

Outline of One-Step *In Vivo* Assembly of Multiple DNA Fragments and Genomic Integration in *K. phaffii*.

For flexible and easy development of *K. phaffii* strains based on stable integration, a direct method concurrently enabling genetic manipulation and genomic integration is required. The conventional genomic integration method is a cumbersome and time-consuming multistep process that comprises circular plasmid construction (DNA ligation or assembly), *E. coli* transformation, plasmid extraction, and linearization of the circular plasmid,[21,22] before introducing linear DNA into *K. phaffii* (Figure 1). Although OE-PCR has been usually utilized to prepare a linear DNA fragment fused *in vitro*,[27,28] this technique may be unsuitable for assembling high-order multiple DNA fragments (*e.g.*, over 10 fragments) due to difficulties in specifically amplifying properly fused DNA fragments, as their number increases.

Here, we established a one-step *in vivo* assembly of multiple DNA fragments and genomic integration method for *K. phaffii*, as a novel technical platform (Figure 1). This system enables an efficient *in vivo* assembly of multiple DNA fragments concurrently with a precise integration of the assembled DNA into a target genomic locus. To make this achievable, we used the Dnl4p-defective strain that exhibits higher HR efficiency than the wild-type strain, owing to decay of the NHEJ activity. Multiple DNA fragments that have homologous sequences correspondent to either mutual fragments (46−50 bp) or the target genomic loci (over 200 bp; 238 and 362 bp) at each end

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

were used to achieve simultaneous DNA assembly and genomic integration. The length of the homologous sequence was set to over 30 bp, referring to the experimental results using the autonomous replicating plasmid vector as a backbone vector (Figure S1) and previous double-crossover recombination study (using 50 bp flanking homologies) in the Dnl4p-defective strain.[34]

Transformation Efficiency in the Wild-Type and the Dnl4p-Defective Strains. Two *K. phaffii* strains, CBS7435 (wild-type strain) and CBS7435 *dnl4 his4* (Dnl4p-defective strain),[34] (Table 1) were used as hosts to test the feasibility of

Table 1. Yeast Strains Used in This Study

name	specific features	source
	strains	
CBS7435	*Komagataella phaffii* wild-type strain (NRRL Y-11430 or ATCC 76273)	ATCC
CBS7435 *dnl4 his4*	*dnl4*Δ *his4*Δ*::ADE1* (Dnl4p-defective strain)	previous study[34]
CBS7435 *ku70*	*ku70*Δ (Ku70p-defective strain)	previous study[34]
CBS7435 *dnl4 ku70 his4*	*dnl4*Δ *ku70*Δ *his4*Δ*::ADE1* (double-defective strain)	previous study[34]

the one-step *in vivo* DNA assembly and single-crossover genomic integration method in *K. phaffii*. A previously constructed plasmid, harboring the expressed *GFP* gene (pPGP_EGFP), was used as a template to prepare DNA fragments for model assembly and genomic integration into the 3′-untranslated region of the *CCA38473* gene (KpT_{38473})[34] (Table 2).

Three DNA fragments were PCR amplified, using the plasmid pPGP_EGFP as a template. In particular, DNA fragments were named Fragment G1, G2, or G3, depending on their specific features (Figure 2A). Fragment G1 contained the *GAPDH* promoter derived from *K. phaffii* (KpP_{GAPDH}) and the *EGFP* gene, Fragment G2 had the *AOX1* terminator derived from *K. phaffii* (KpT_{AOX1}) and KpT_{38473} (1−238), and Fragment G3 harbored KpT_{38473} (239−600) and $G418^R$ (Table 2). Both the wild-type strain and the Dnl4p-defective strain were transformed with the mixture of Fragment G1 (0.3 pmol), Fragment G2 (0.3 pmol), and Fragment G3 (0.015 pmol) in various combinations using electroporation (Figure 2B). The ratio of the backbone vector to other DNA fragments was set to over 1:10, referring to the experimental results using the autonomous replicating plasmid vector as a backbone vector in *K. phaffii* (Figure S1).

To investigate transformation efficiency, colony-forming units (cfu) were determined with the reference to the DNA Fragment G3 that included the G418-resistance marker gene, $G418^R$. As a result, even when Fragment G1 and Fragment G2 were not used, which are necessary for proper assembly and homologous recombination, the wild-type CBS7435 strain yielded a relatively large number of colonies on a selective plate (transformation efficiency = $1.2 \pm 0.3 \times 10^3$ cfu/μg) (Figure 2B). It may be hypothesized that DNA fragments were non-specifically assembled and/or, consequently, randomly integrated into off-target genetic loci under the influence of the NHEJ mechanism in the wild-type strain. Interestingly, for the assembly that included the Dnl4p-defective strain as a host, combining all fragments (Fragments G1 + G2 + G3) yielded outnumbered colonies (transformation efficiency = $5.9 \pm 2.0 \times 10^3$ cfu/μg) (Figure 2B,C), whereas the exclusion of all

Table 2. Vectors and Fragments Used in This Study

name	specific features	source
	template plasmids	
pUC19	cloning vector	Takara Bio
pPGP_EGFP	KpP_{GAPDH}_*EGFP*_KpT_{AOX1} expression cassette, KpT_{38473} homologous sequence, and $G418^R$ in pUC19	previous study[34]
	fragments for the GFP experiment	
Fragment G1	KpP_{GAPDH}_*EGFP*	this study
Fragment G2	KpT_{AOX1} and KpT_{38473} (1−238)	this study
Fragment G3	KpT_{38473} (239−600) and $G418^R$ in pUC19	this study
	fragments for the lycopene experiment	
Fragment L1	KpP_{GAPDH}_*PaCrtE*_ScT_{RPL3}	this study
Fragment L2	$KpP_{CCA36283}$_*PaCrtB*_ScT_{PRC1}	this study
Fragment L3	KpP_{TEF2}_*PaCrtI*_ScT_{EFM1} and KpT_{38473} (1−238)	this study
Fragment L4	KpT_{38473} (239−600) and $G418^R$ in pUC19	this study
Fragment L1′	ScP_{TEF1}_*PaCrtE*_ScT_{EFM1}	this study
Fragment L2′	ScP_{TDH1}_*PaCrtB*_ScT_{PRC1}	this study
Fragment L3′	ScP_{TEF2}_*PaCrtI*_ScT_{RPL3}_KpT_{38473} (1−238)	this study
Fragment L4	KpT_{38473} (239−600) and $G418^R$ in pUC19	this study
Fragment DC1	KpP_{AOX1}_ScP_{TEF1}	this study
Fragment DC2	*PaCrtE*	this study
Fragment DC3	ScT_{EFM1}	this study
Fragment DC4	ScP_{TDH1}	this study
Fragment DC5	*PaCrtB*	this study
Fragment DC6	ScT_{PRC1}	this study
Fragment DC7	ScP_{TEF2}	this study
Fragment DC8	*PaCrtI*	this study
Fragment DC9	ScT_{RPL3}	this study
Fragment DC10	$G418^R$_KpT_{AOX1}	this study

fragments resulted in almost no colonies on the selective plates (Figure 2B). These results imply that disruption of the Dnl4p gene is able to suppress the NHEJ activity, thereby greatly decreasing non-specific assembly of DNA fragments and their random genomic integration. This striking effect is the aftermath of the gene targeting efficiency obtained in our previous study on the Dnl4p-defective strain.[34] Similarly, the Dnl4p deletion would have helped the definitive appearance of accurately assembled fragments by inhibiting the DNA end-joining activity.

Confirmation of Accurate Assembly and Integration in the Dnl4p-Defective Strain. To test whether the appeared colonies have accurately assembled the fragments and whether the latter has been integrated into the target loci, the GFP expression of 24 transformants obtained from *in vivo* assembly with all the three fragments (Fragments G1 + G2 + G3) was evaluated first. After incubating randomly selected colonies on selective agar plates, GFP fluorescence was detected using a fluorescent image analyzer (Figure 3A). As a result of the experiments in which the wild-type strain was used as a host, 17 out of 24 transformants (71%) exhibited fluorescence. Seven non-fluorescent transformants had probably unsuccessfully integrated Fragment G3 (vector backbone harboring the G418-resistance gene, $G418^R$) or the fusion occurred between Fragment G2 and Fragment G3, omitting Fragment G1 (GFP expression cassette). In contrast, when the

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

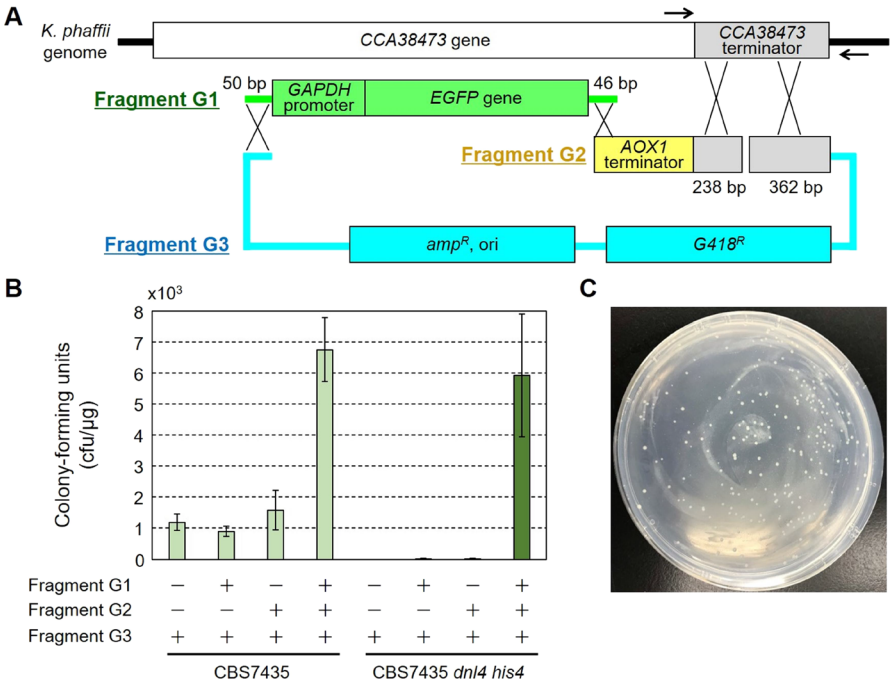


Figure 2. One-step *in vivo* assembly of DNA fragments and genomic integration of GFP expression components in *K. phaffii*, using three DNA fragments. (A) DNA fragments used for GFP assembly and integration experiment. The DNA fragment G1 contains the KpP_{GAPDH} and the *EGFP* gene. The DNA fragment G2 contains the KpT_{AOX1} and KpT_{38473} (1−238). The DNA fragment G3 contains KpT_{38473} (239−600) and $G418^R$. Black arrows indicate the positions of the primers #7 and #8 used for confirmation of the genomic integration. (B) Transformation efficiency with various combinations of the three DNA fragments. The number of colony-forming units (cfu) per microgram of the Fragment G3 DNA was determined by transformation of the CBS7435 wild-type and the *dnl4 his4* yeast strains. The cfu value was defined as transformation efficiency. Standard deviation is calculated according to the number of cfu obtained in triplicates (n = 3). (C) Representative image of colonies formed by Dnl4p-defective transformants that contained Fragments G1 + G2 + G3.

Dnl4p-defective strain was used, all 24 transformants (100%) showed fluorescence, assuming that all the three fragments have been accurately assembled and integrated in these transformants.

In the subsequent experiments, both accurate assembly and integration were confirmed by PCR amplification at the target locus (*CCA38473* terminator region) (Figure 2A). PCR products were electrophoretically separated on agarose gels to check whether the resulting DNA bands corresponded to the expected size, *i.e.*, 6725 bp. The wild-type strain generated 15 out of 24 transformants (63%) that had bands of the predicted size (Figure 3B). Although two transformants from the wild-type strain (sample nos. 7 and 20) did not present bands of the anticipated size, they exhibited GFP fluorescence (Figure 3A). This suggests that the GFP expression cassette (Fragment G1) and the $G418^R$ vector backbone (Fragment G3) were integrated in an unexpected genetic locus in these two transformants. Moreover, 3 out of 24 transformants (sample nos. 7 and 20—fluorescent and sample no. 23—non-fluorescent) never generated PCR products, while 6 out of 24 transformants (sample nos. 2, 5, 14, 16, 17, and 19—all being non-fluorescent) provided PCR products, but the band sizes were shorter than expected (Figure 3B). In contrast, the Dnl4p-defective strain covered all 24 transformants (100%) with bands of the correctly resulting sizes (Figure 3B), indicating that the loss of NHEJ activity by the deletion of Dnl4p drastically increased both the efficiency of accurate assembly of multiple DNA fragments and genomic integration in *K. phaffii*.

At last, to grasp more detailed genomic information as whether the DNA vectors were accurately assembled, each amplified DNA sequence from the target loci was confirmed by Sanger sequencing analysis (Table 3). In the wild-type strain, the alignment of the PCR-amplified products and sequencing data suggested that 6 out of 24 transformants (displaying shorter bands, Table 3) non-specifically combined only Fragment G2 and Fragment G3, by omitting Fragment G1, and integrated this combination into the target *CCA38473* terminator locus owing the activity of the NHEJ machinery (Figure S2). As described above, these 6 transformants certainly showed no GFP fluorescence (Figure 3A). In addition, 2 out of 24 transformants (Table 3) were found to have a mutation in the *GAP* promoter sequence (sample nos. 9 and 10) and in the overlap region of Fragments G1 and G2 (sample no. 9) (data not shown). The remaining 13 transformants (54%) showed the correct sequences. In contrast, in the Dnl4p-defective strain, all 24 transformants (100%) showed almost completely correct sequences (Table 3). Unexpectedly, 3 out of 24 transformants (sample nos. 8, 12, and 17 from the Dnl4p-defective strain) showed an additional mononucleotide $(A)_1$ in the repetitive $(A)_9$ sequence in Fragment G2 (Table 3 and Figure S3). However, this repetitive $(A)_9$ sequence is located in the middle of the *AOX1* terminator and distantly from the homologous regions, which may point out to an error caused by the mononucleotide slippage during the PCR reaction[36] and not by the *in vivo* DNA assembly. Therefore, the efficiency of the method composed of *in vivo* DNA assembly and genomic integration in the Dnl4p-defective strain can be estimated to reach 100%. Thus, we demonstrated that the NHEJ-defective *dnl4* strain enabled one-step *in vivo* assembly of multiple DNA fragments and genomic integration in *K. phaffii*.

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

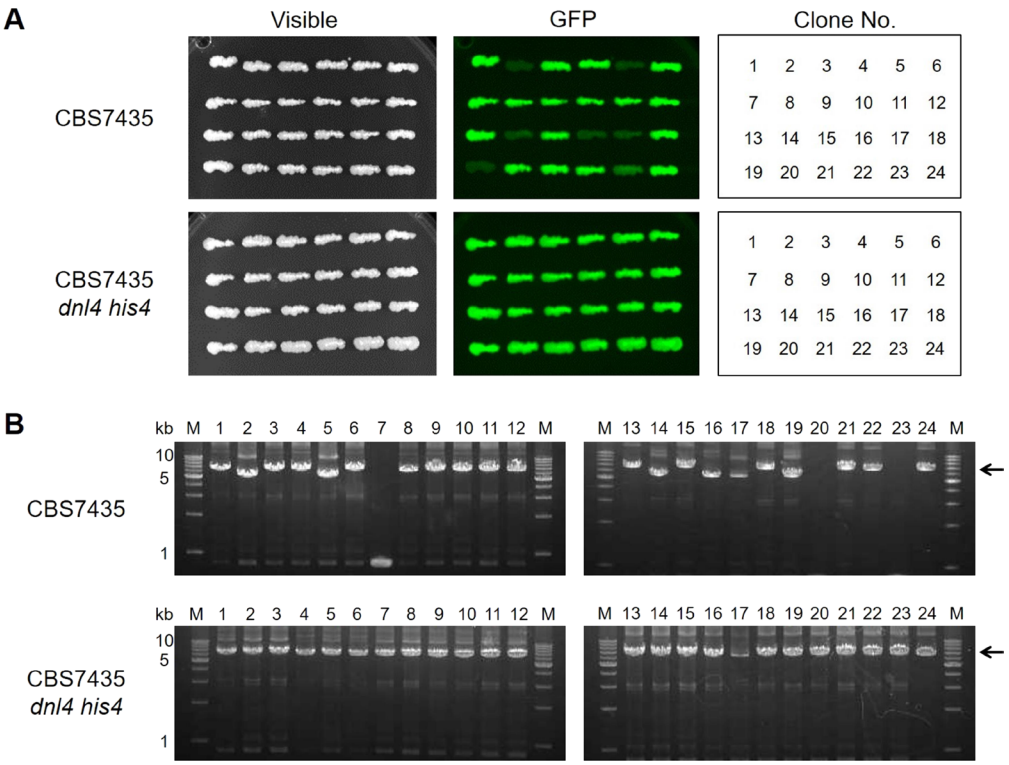


Figure 3. Evaluation of the accuracy of simultaneous *in vivo* assembly of DNA fragments and genomic integration. Performance of GFP assembly and integration was tested using the CBS7435 wild-type and the *dnl4 his4* yeast strains. (A) Evaluation of gene functionality in transformants by the GFP fluorescence. The 24 transformants obtained by transformation with Fragments G1 + G2 + G3 were randomly selected and inoculated on selective agar plates. After 3 day incubation, GFP fluorescence was observed using a fluorescent image analyzer. (B) Confirmation of genomic integration into the target locus (*CCA38473* terminator region) by PCR amplification. Genomic DNA was extracted from each out of 24 transformants, and genomic integration was confirmed by PCR using the primers #7 and #8 (Figure 2A). PCR products were electrophoretically separated, and 1 kb DNA ladder (M) was used to test the expected product size of 6725 bp (black arrows).

Table 3. PCR Amplification and Sequencing Analysis of the Target Genomic Loci for Checking the Efficiency of One-Step *In Vivo* Assembly of DNA Fragments and Genomic Integration by the GFP Experiment

strain	not amplified[a]	short[a]	mutation[a]	correct[a]
CBS7435	3	6 (1)[b]	2	13 (1)[b]
CBS7435 *dnl4 his4*	0	0	0	24 (3)[b]

[a]The categorized number of transformants obtained by PCR amplification and Sanger sequencing analysis is displayed. The target *CCA38473* terminator locus was PCR amplified using the primers #9 and #10. PCR products were sequenced using the primers #9, #10, and #11. [b]In the Fragment G2 of 2 transformants from the wild-type strain and 3 transformants from the Dnl4p-defective strain, additive mononucleotide $(A)_1$ was inserted into the repetitive $(A)_9$ sequences. This was probably due to the mononucleotide slippage during the PCR reaction.

Application of the Developed Method to Metabolic Pathway Genes Included in Lycopene Biosynthesis. Next, the established method was applied to the assembly and integration of three gene expression cassettes for enzymes included in the lycopene biosynthesis pathway (CrtE, CrtB, and CrtI derived from *Pantoea ananatis*[37]).

Initially, three expression cassettes, harboring different sets of native promoters of *K. phaffii* (KpP_{GAPDH}, $KpP_{CCA36283}$, or KpP_{TEF2}) and heterologous terminators of *S. cerevisiae* (ScT_{RPL3}, ScT_{PRC1}, or ScT_{EFM1}), were prepared and accordingly named Fragments L1, L2, and L3, respectively (Figure 4A and Table 2). All the DNA fragments contained homologous sequences correspondent to either mutual fragments (50 bp) or the target genomic locus (238 and 362 bp, the same as for the GFP experiment) at each end (Figure 4A). The DNA fragment that contained KpT_{38473} (239−600) and $G418^R$, named Fragment L4, had exactly the same sequence as Fragment G3 (Figure 2A), and it was used as the vector backbone (Figure 4A). The Dnl4p-defective strain, representing a host, was transformed with the mixture of Fragment L1 (0.2 pmol), Fragment L2 (0.2 pmol), Fragment L3 (0.2 pmol), and Fragment L4 (0.01 pmol) by electroporation. After long-term incubation on selective agar plates (for 8 days), colonies were obtained (transformation efficiency = $9.3 \pm 0.3 \times 10^4$ cfu/μg) (Figure 4B). This delayed colony formation might be the result of an overload in either protein synthesis or lycopene biosynthesis. After culturing four pigmented colonies in liquid medium, the cells displayed clear pink colors, although each clone showed slightly different pigmentation (Figure 4E). The result of average copy number analysis showed that approximately a single copy of the *crtE*, *crtB*, and *crtI* gene was detected per cell (Figure S4). Since the gene copy numbers did not show variations, this different pigmentation might be from the difference of protein expression. Targeting efficiency into the target genomic locus (*CCA38473* terminator region) was examined among the randomly selected pigment colonies, but the integration was detected in only 15 out of 35 transformants (43%) (Figure 4C,D). Therefore, to increase targeting efficiency, longer homologous sequences (100 or 150 bp) were used as homology arms for connecting mutual fragments. However, longer homologous sequences barely improved

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

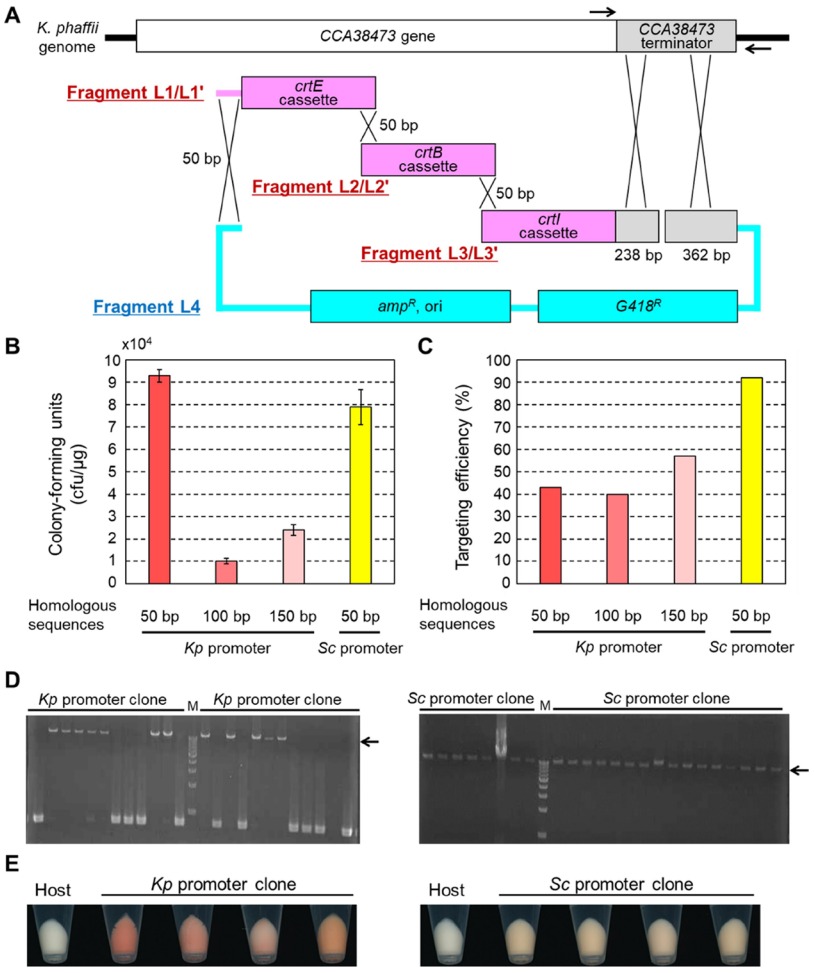


Figure 4. One-step *in vivo* assembly of DNA fragments and integration for reconstitution of the lycopene biosynthesis pathway in *K. phaffii* using four DNA fragments. (A) DNA fragments used for assembly and integration of the lycopene production component. The DNA fragment L1/L1′ contained the *crtE* gene, L2/L2′ contained the *crtB* gene, and the DNA fragment L3/L3′ contained the *crtI* gene and KpT_{38473} (1−238). The DNA fragment L4 (blue) was the same as the DNA fragment G3 (Figure 2A) that contained KpT_{38473} (239−600) and $G418^R$. These four DNA fragments were prepared by PCR and used for the experiment. Black arrows indicate the positions of the primers #7 and #8 used for confirmation of the genomic integration. (B) Transformation efficiency obtained using *K. phaffii* and *S. cerevisiae* promoters as the components of the DNA fragments. Various lengths of homologous sequences were also tested. The number of colony-forming units (cfu) per microgram of the Fragment L4 DNA was determined by transformation of the CBS7435 *dnl4 his4* yeast strain. The cfu value was defined as transformation efficiency. Standard deviation is calculated according to the number of cfu recorded in triplicate (n = 3). (C) Targeting efficiency into the target genomic loci (*CCA38473* terminator region). Genomic DNA was extracted from each transformant and genomic integration was confirmed by PCR using the primers #7 and #8. Targeting efficiency was calculated as the ratio of the number of correct transformants to that of all transformants tested. (D) Confirmation of genomic integration by PCR amplification. PCR products were electrophoretically separated along with 1 kb DNA ladder (M) to test matching the anticipated size (black arrows). (E) Representative image of cells pigmented by cultivation in the YPD + G418 liquid medium. The clones represent transformants that contained either Fragments L1 + L2 + L3 + L4 (*Kp* promoter) or Fragments L1′ + L2′ + L3′ + L4 (*Sc* promoter) in the Dnl4p-defective strain.

targeting efficiency (14 out of 35 transformants (40%), using 100 bp homologous sequences and 20 out of 35 transformants (57%), using 150 bp homologous sequences) (Figure 4C and Figure S5) and adversely rather decreased transformation efficiency ($1.0 \pm 0.1 \times 10^4$ and $2.4 \pm 0.2 \times 10^4$ cfu/μg, using 100 and 150 bp homologous sequences, respectively) (Figure 4B). We hypothesize that a low targeting efficiency, observed even for the Dnl4p-defective strain, has been caused by using promoter sequences cognate to *K. phaffii*, thus generating non-specific integrations in long sequences (~1 kb), having the exact matches with the host genome.

Then, we replaced the *K. phaffii* promoters (KpP_{GAPDH}, $KpP_{CCA36283}$, and KpP_{TEF2}) with the *S. cerevisiae* promoters (ScP_{TEF1}, ScP_{TDH1}, and ScP_{TEF2}) that have been confirmed to function in *K. phaffii* cells (Figure S6) for each expression cassette (Figure 4A and Table 2). Here, 50 bp long homologous sequences were used for assembling mutual fragments. Other conditions were copied from the previous experiment on lycopene assembly and integration, except for the preparation of the fragment mixture: Fragment L1′ (0.3 pmol), Fragment L2′ (0.3 pmol), Fragment L3′ (0.3 pmol), and Fragment L4 (0.015 pmol). After 3 days incubation on selective agar plates, colonies were usually obtained, and transformation efficiency was similar to that of the *K. phaffii* promoters ($7.9 \pm 0.8 \times 10^4$ cfu/μg) (Figure 4B). Although it was difficult to discriminate between the pigment colors of colonies from photography images, virtually, all colonies grown on selective plates displayed faint but noticeable light pink

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

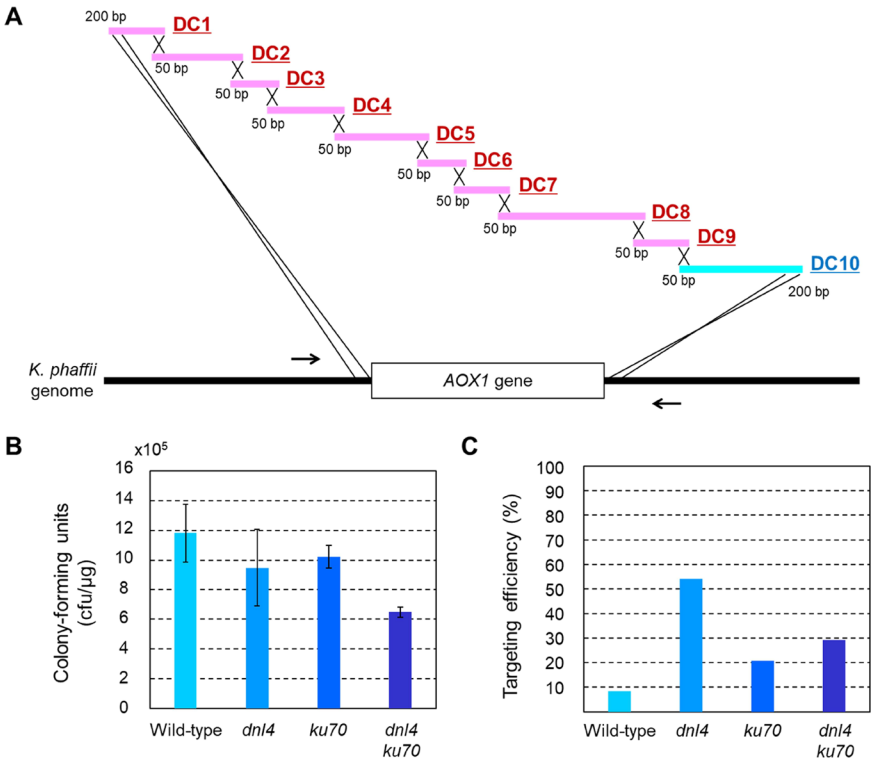


Figure 5. One-step *in vivo* assembly of DNA fragments and integration for reconstitution of the lycopene biosynthesis pathway in *K. phaffii*, using 10 DNA fragments. (A) DNA fragments used for assembly and integration of lycopene production component. The DNA fragment DC1 to DC3 contained KpP_{AOX1} and the *crtE* cassette, DC4 to DC6 contained the *crtB* cassette, and the DNA fragment DC7 to DC9 contained the *crtI* cassette. The DNA fragment DC10 contained $G418^R$ and KpT_{AOX1}. These 10 DNA fragments were prepared by PCR and used for the experiment. Black arrows indicate the positions of the primers #54 and #55 used for confirmation of the genomic integration. (B) Transformation efficiency obtained using each strain. The number of colony-forming units (cfu) per microgram of the Fragment DC10 DNA was determined by transformation of each strain. The cfu value was defined as transformation efficiency. Standard deviation is calculated according to the number of cfu recorded in triplicate ($n = 3$). (C) Targeting efficiency into the target genomic loci (*AOX1* gene region). Genomic DNA was extracted from each transformant and genomic integration was confirmed by PCR, using the primers #54 and #55. Targeting efficiency was calculated as the ratio of the number of correct transformants to that of all transformants tested.

coloration (data not shown). Moreover, the cells cultured in a liquid medium also displayed weak but obvious pink pigmentation (Figure 4E). This was the result of the preferential selection of the *S. cerevisiae* promoters over the *K. phaffii* ones that possessed a weaker transcription activity. To solve this problem, employing either heterologous or synthetic promoters[38] might be needed to enhance transcription in *K. phaffii*. However, as expected, the result of genomic integration analysis showed that 22 out of 24 transformants (92%) had the exactly sized bands, indicating accurate DNA assembly and genomic integration into the correct targeting locus (Figure 4C,D). The remaining 2 transformants might have redundantly assembled any fragment or have tandemly integrated the assembled full-length fragments (Figure 4D). Thus, we successfully proved that the technique reported in this study could be applied to both assembly and integration of expression cassettes for lycopene biosynthesis pathway enzymes, which needed longer size and higher number of DNA fragments than in the GFP experiment.

Finally, to examine whether the established method was useful not only for single-crossover integration but also for double-crossover integration, we tested concurrent high-order multiple DNA assembly and genomic integration (replacement) for the gene knock-out of *AOX1* by using 10 DNA fragments. To further examine whether the combination of NHEJ-related proteins affected the targeting efficiency, we evaluated Ku70p- and Dnl4p/Ku70p double-defective strains as well as the Dnl4p-defective strain. Three expression cassettes for the lycopene biosynthesis pathway enzymes were separated to 9 DNA fragments, named Fragment DC1 to DC9 (Figure 5A and Table 2). The DNA fragment that contained $G418^R$, named Fragment DC10, was used as the selectable marker (Figure 5A and Table 2). All the DNA fragments contained homologous sequences correspondent to either mutual fragments (50 bp) or the target genomic locus (200 bp; *AOX1* promoter and terminator, KpP_{AOX1} and KpT_{AOX1}, respectively) at each end (Figure 5A). The wild-type, Dnl4p-, Ku70p-, and Dnl4p/Ku70p double-defective strains (wild-type, *dnl4*, *ku70*, and *dnl4 ku70* strains, respectively) were transformed with the mixture of Fragment DC1 to DC9 (0.2 pmol) and Fragment DC10 (0.05 pmol) by electroporation (Table 1). Colonies were obtained after 3 days incubation on selective agar plates, and the transformation efficiency was sufficient (Figure 5B). Expectedly, the result of genomic integration analysis showed that 13 out of 24 transformants (54%) from the Dnl4p-defective strain had bands of the expected size (Figure 5C and Figure S7). However, unexpectedly, the Dnl4p/Ku70p double-defective strain did not improve the targeting efficiency (7 out of 24 transformants (29%)) (Figure 5C and Figure S7). We hypothesize that deletion of Ku70p might lead a trade-off relationship that suppresses the NHEJ activity and facilitates

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

the degradation of DNA fragments, because Ku70p plays an important role in protecting the ends of DNA. Therefore, Ku70p overexpression may be effective for improving targeting efficiency.

In conclusion, we demonstrated that the developed one-step method enables concurrent *in vivo* DNA assembly and its integration into the target genetic locus of *K. phaffii*. The use of a Dnl4p-defective strain, hindering employment of the NHEJ mechanism, increased the efficiency of accurate assembly of multiple DNA fragments and precise target integration. For assembling mutual DNA fragments, 50 bp long homologous sequences were proved to be large enough. Although a suitable length for the concurrent genomic integration was not determined, 200 bp long homologous regions allowed for efficient and precise integration into the target genetic locus. The use of native, endogenous *K. phaffii* promoters led to a low targeting efficiency and was therefore unfavorable, probably due to the generation of non-specific integrations in long, exactly matched sequences in the host genome. While we presented one-step assembly and integration using 10 fragments, the limits of number and length of the *K. phaffii* DNA fragments used for the development of this method remained unexamined in this study, anticipating that further investigation would disclose these detailed features in the near future. By combining this method with *RAD52* overexpression, *MPH1* deletion, and/or the CRISPR-Cas9 system,[35] it would be possible to simultaneously integrate dozens of DNA fragments into multiple genomic loci in *K. phaffii*, as well as the CasEMBLR system in *S. cerevisiae*.[9] The technique described here will surely help to considerably decrease both labor and time for the construction of *K. phaffii* strains and will also enable diverse and more flexible vector construction in *K. phaffii*, thus representing a powerful tool for synthetic biology.

METHODS

Yeast Strains and Media. All strains used in this study are listed in Table 1. The *K. phaffii* strain CBS7435 (NRRL Y-11430 or ATCC 76273) was obtained from the ATCC (American Type Culture Collection; Manassas, VA, USA). The CBS7435 *dnl4 his4* strain, the CBS7435 *ku70* strain, and the CBS7435 *dnl4 ku70 his4* strain used in this study were constructed previously.[34] The YPD medium contained 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. The YNB solution contained 1.7 g/L yeast nitrogen base (YNB) without amino acids or ammonium sulfate (BD Biosciences) and 1 g/L monosodium glutamate (Nacalai Tesque). The SD agar plate contained 1.7 g/L YNB without amino acids or ammonium sulfate (BD Biosciences), 1 g/L monosodium glutamate, 20 g/L glucose, and 15 g/L agar.[39] The YPD agar plate contained the YPD medium and 20 g/L agar. Where appropriate, 40 mg/L histidine (FUJIFILM Wako Pure Chemical, Osaka, Japan) and 500 mg/L G418 (Nacalai Tesque) were added.

Plasmid Construction and Preparation of DNA Fragments. All vectors, plasmids, and fragments used in this study are listed in Table 2. The primers used for vector and strain construction are listed in Table S1. The plasmid pPGP_EGFP, used to prepare the DNA fragments for assembly (Fragments G1, G2, and G3; Table 2), was derived from pUC19 (Takara Bio, Shiga, Japan) and was previously constructed as a genomic integration plasmid.[34] The plasmid pPGP_EGFP contained the *GAPDH* promoter (P_{GAPDH}) derived from *K. phaffii* (KpP_{GAPDH}), the enhanced green fluorescent protein (*EGFP*) gene, the *AOX1* terminator (T_{AOX1}) derived from *K. phaffii* (KpT_{AOX1}), 3′-untranslated (terminator) region of the *CCA38473* gene (T_{38473}) on the *K. phaffii* genome (KpT_{38473}), and the G418-resistance marker ($G418^R$).

A DNA fragment that contained KpP_{GAPDH}_*EGFP*, a fragment with KpT_{AOX1} and KpT_{38473} (1−238), and the pUC19 fragment with KpT_{38473} (239−600) and $G418^R$ were PCR amplified using pPGP_EGFP as a template and the primer pairs: #1/#2 (for Fragment G1, 1.3 kb), #3/#4 (for Fragment G2, 0.6 kb), or #5/#6 (for Fragment G3, 4.1 kb), respectively. Each fragment for the GFP (EGFP) expression component was designed to have homologous recombination sequences for assembly and integration at both ends (46−50 bp long for connection between each DNA fragment and 238/362 bp long for integration into the *K. phaffii* genome) (Figure 2A and Table S2).

The genes encoding for CrtE, CrtB, and CrtI open reading frames (ORFs) in the lycopene metabolic pathway, derived from *Pantoea ananatis*,[37] were synthesized using GeneArt Gene Synthesis (Thermo Fisher Scientific, Waltham, MA, USA). Each DNA fragment for promoters (native *GAPDH*, *CCA36283*, and *TEF2* promoters of *K. phaffii* (KpP_{GAPDH}, $KpP_{CCA36283}$, and KpP_{TEF2}) or heterologous *TEF1*, *TDH1*, and *TEF2* promoters of *S. cerevisiae* (ScP_{TEF1}, ScP_{TDH1}, and ScP_{TEF2})) was obtained by PCR amplification from the genome of the *K. phaffii* CBS7435 and the *S. cerevisiae* BY4741 strains. DNA fragments of each promoter and the respective CrtE, CrtB, or CrtI ORF were assembled into the BamHI/SpeI and SpeI/XhoI sites of three types of terminator plasmids based on pPGP_EGFP, containing heterologous *RPL3*, *PRC1*, and *EFM1* terminators of *S. cerevisiae* (ScT_{RPL3}, ScT_{PRC1}, and ScT_{EFM1})[40] by In-Fusion technology (Takara Bio). The resultant plasmids, carrying CrtE, CrtB, or CrtI expression cassette were named pCrtE, pCrtB, and pCrtI (*K. phaffii* promoters: KpP_{GAPDH}, $KpP_{CCA36283}$, and KpP_{TEF2}) or pCrtE′, pCrtB′, and pCrtI′ (*S. cerevisiae* promoters: ScP_{TEF1}, ScP_{TDH1}, and ScP_{TEF2}), respectively. Three DNA fragments of CrtE, CrtB, and CrtI expression cassettes driven by the *K. phaffii* promoters were PCR amplified using template plasmids (pCrtE, pCrtB, or pCrtI) and tandemly connected into the BamHI/BglII sites of the plasmid pPGP_EGFP, named pCrtEBI, using the In-Fusion technology (Takara Bio). The CrtE, CrtB, and CrtI expression cassettes, carrying homologous regions of different lengths at both ends (50, 100, or 150 bp), were prepared as described below.

Three expression cassettes for CrtE, CrtB, and CrtI (lycopene biosynthesis pathway enzymes) were PCR amplified using pCrtEBI (*K. phaffii* promoters) or pCrtE′, pCrtB′, and pCrtI′ (*S. cerevisiae* promoters) and primer pairs (described below) for preparation of each of the three DNA fragments (Fragments L1, L2, and L3 or L1′, L2′, and L3′, respectively) (Table 2). The amplified fragments were of 2.3, 2.3, and 3.2 kb or 1.8, 2.2, and 2.8 kb long, respectively (Table S2). A DNA fragment that contained KpT_{38473} (239−600) and $G418^R$ (4.1 kb) was PCR amplified by the same procedure and using the same primers as for Fragment G3, and the resulting fragment was named Fragment L4. Each fragment contained homologous sequences for assembling mutual fragments (the lengths are described below) or the target genetic locus (238 and 362 bp, the same as for GFP) at each end (Figure 4A). When expression cassettes were combined with the *K. phaffii*

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

promoters, homologous sequences of different lengths (50, 100, or 150 bp) were prepared for assembly of mutual fragments using the corresponding primer pairs (#12/#13, #22/#23, or #24/#25 for Fragment L1; #14/#15, #26/#27, or #28/#29 for Fragment L2; and #16/#4, #30/#4, or #31/#4 for Fragment L3, respectively) and pCrtEBI as a template. When expression cassettes were combined with the *S. cerevisiae* promoters, only 50 bp of the homologous sequences were used to assemble mutual fragments using the corresponding primer pairs (#17/#18 for Fragment L1′, #19/#20 for Fragment L2′, and #21/#4 for Fragment L3′).

Three expression cassettes for double-crossover integration were PCR amplified using pCrtE′, pCrtB′, and pCrtI′ (*S. cerevisiae* promoters and terminators) and primer pairs (described below) for preparation of each of the 9 DNA fragments (Fragments DC1, DC2, DC3, DC4, DC5, DC6, DC7, DC8, and DC9, respectively) (Table 2 and Table S2). A DNA fragment that contained $G418^R$ and KpT_{AOX1} was PCR amplified using pPGP_EGFP and named Fragment DC10 (Table 2 and Table S2). Each fragment contained homologous sequences (50 bp) for assembling mutual fragments or the target genetic locus (200 bp) at each end (Figure 5A and Table S2). The DNA fragments were prepared using the corresponding primer pairs (#32/#33/#34 for Fragment DC1, #35/#36 for Fragment DC2, #37/#38 for Fragment DC3, #39/#40 for Fragment DC4, #41/#42 for Fragment DC5, #43/#44 for Fragment DC6, #45/#46 for Fragment DC7, #47/#48 for Fragment DC8, #49/#50 for Fragment DC9, and #51/#52/#53 for Fragment DC10, respectively).

***In Vivo* DNA Assembly and Genomic Integration.** The *K. phaffii* CBS7435 strain, CBS7435 *dnl4 his4* strain, CBS7435 *ku70* strain, and CBS7435 *dnl4 ku70 his4* strain were transformed by electroporation as previously described.[34] In short, competent cells (60 μL) were mixed with DNA fragments in a 0.2 cm electroporation cuvette, incubated on ice for 5 min, and electroporated in a Bio-Rad Gene Pulser (Bio-Rad Laboratories, Hercules, CA, USA) with the following settings: 1500 V, 25 μF, and 200 ohms. After pulsing, 1 mL of YPD medium was added immediately to the cuvette, and the cells were transferred to a sterile tube.

For the GFP expression experiment, the tube was subsequently incubated at 30 °C for 1 h without shaking. Thereafter, the cells were harvested by centrifugation at 3000 × *g* at 25 °C for 5 min, and all supernatants were discarded. The cells were resuspended in 1 mL YNB solution and harvested by centrifugation at 3000 × *g* at 25 °C for 5 min, after which the 950 μL supernatant was discarded. The cells were resuspended in the remaining YNB supernatant (about 50 μL) and plated onto either SD + G418 (containing 500 mg/L G418; for the CBS7435 wild-type strain) or SD + G418 + His (containing 500 mg/L G418 and 40 mg/L histidine; for the CBS7435 *dnl4 his4* strain) selective agar plates.[39]

For the lycopene production experiment using single-crossover integration (Figure 4A), the CBS7435 *dnl4 his4* strain, representing a host, was used to prepare competent cells. For the lycopene production experiment using double-crossover integration (Figure 5A), the CBS7435 strain, the CBS7435 *dnl4 his4* strain, the CBS7435 *ku70* strain, and the CBS7435 *dnl4 ku70 his4* strain were used to prepare competent cells. After pulsing and adding 1 mL of YPD medium, the tube was incubated at 30 °C for 2 h without shaking. Thereafter, 100 μL of suspension (single-crossover integration) or 10 μL of suspension (double-crossover integration) was plated onto YPD + G418 selective agar plates.

The agar plates were incubated at 30 °C for either 3 days (for the GFP expression experiment and for the lycopene production experiment using the *S. cerevisiae* promoters) or 8 days (for the lycopene production experiment using the *K. phaffii* promoters). The number of colony-forming units (cfu) per microgram (μg) of a vector backbone (Fragment G3, Fragment L4, or Fragment DC10) was determined from the number of colonies generated on the selective agar plates and defined as transformation efficiency.

Evaluation of GFP Fluorescence and Lycopene Pigmentation. After the GFP expression component assembly, 24 transformants were randomly sampled from across the three independent selective plates and replicated onto either SD + G418 or SD + G418 + His selective agar plates. After incubation for 3 days, green fluorescence of cells was recorded on ImageQuant LAS 4000 (GE Healthcare, Chicago, IL, USA), using λ_{ex} = 460 nm and the detection filter 510DF10 (510 nm ± 5 nm).

After the lycopene production component assembly, its pigmentation was investigated with reference to our previous paper.[41] In brief, colored transformants were selected and replicated onto the YPD + G418 selective agar plate, followed by cultivation in the liquid YPD medium at 30 °C for 48 h. After washing off the yeast cells twice with distilled water, photo images of yeast cell pellets were captured using a scanner.

Confirmation of Genomic Integration and Sequencing Analysis. Genomic DNA was extracted from cell suspensions using Kaneka Easy DNA Extraction Kit version 2 (Kaneka, Osaka, Japan). After centrifugation of cell extracts, the supernatant was used as a template for PCR. Genomic integration into the target locus (*CCA38473* terminator) was confirmed by PCR, using KOD Fx Neo DNA polymerase (TOYOBO, Osaka, Japan) and the primers #7 and #8 that recognized sequences both upstream and downstream of the integrated loci (Figures 2A and 4A). Genomic integration into the target locus (*AOX1* gene) was confirmed by PCR, using KOD Fx Neo DNA polymerase (TOYOBO) and the primers #54 and #55 that recognized sequences both upstream and downstream of the integrated loci (Figure 5A).

Assembly of multiple DNA fragments from the GFP expression components was confirmed by Sanger sequencing analysis. The target locus, including Fragments G1 and G3, was PCR amplified from the extracted DNA using the primers #9 and #10 and KOD One PCR Master Mix (TOYOBO). The PCR products were purified using a Wizard SV 96 PCR Clean-Up System (Promega, Fitchburg, WI, USA) and sequenced by GENEWIZ (South Plainfield, NJ, USA) with appropriate primers (#9, #10, and #11) (Table S1).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00302.

Primer and fragment sequences used in the study, alignment of the sequence data, and electrophoretic confirmation of successful genomic integration (PDF)

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654

AUTHOR INFORMATION

Corresponding Author

Jun Ishii − *Engineering Biology Research Center and Graduate School of Science, Technology and Innovation, Kobe University, Kobe 657-8501, Japan;* orcid.org/0000-0003-2568-515X; Email: junjun@port.kobe-u.ac.jp

Authors

Teruyuki Nishi − *Graduate School of Science, Technology and Innovation, Kobe University, Kobe 657-8501, Japan; Biotechnology Research Laboratories, Pharma & Supplemental Nutrition Solutions Vehicle, Kaneka Corporation, Takasago 676-8688, Japan*

Yoichiro Ito − *Graduate School of Science, Technology and Innovation and Engineering Biology Research Center, Kobe University, Kobe 657-8501, Japan*

Yasuyuki Nakamura − *Graduate School of Science, Technology and Innovation and Engineering Biology Research Center, Kobe University, Kobe 657-8501, Japan;* orcid.org/0000-0002-0093-3754

Taiki Yamaji − *Graduate School of Science, Technology and Innovation, Kobe University, Kobe 657-8501, Japan; Technology Research Association of Highly Efficient Gene Design (TRAHED), Kobe 650-0047, Japan*

Noriko Hashiba − *Technology Research Association of Highly Efficient Gene Design (TRAHED), Kobe 650-0047, Japan*

Masaya Tamai − *Technology Research Association of Highly Efficient Gene Design (TRAHED), Kobe 650-0047, Japan*

Yoshihiko Yasohara − *Biotechnology Research Laboratories, Pharma & Supplemental Nutrition Solutions Vehicle, Kaneka Corporation, Takasago 676-8688, Japan*

Akihiko Kondo − *Graduate School of Science, Technology and Innovation, Engineering Biology Research Center, and Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe University, Kobe 657-8501, Japan; Center for Sustainable Resource Science, RIKEN, Yokohama 230-0045, Japan;* orcid.org/0000-0003-1527-5288

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.1c00302

Author Contributions

T.N., Y.I., Y.N., J.I., and A.K. conceived and designed the experiments. T.N., Y.I., T.Y., N.H., Y.N., and M.T. performed the experiments. T.N., Y.I., Y.N., and J.I. analyzed data. Y.Y., J.I., and A.K. supervised the study. J.I. and A.K. coordinated the research project. T.N., Y.I., Y.N., and J.I. wrote the manuscript. All authors read and approved the manuscript.

Notes

The authors declare the following competing financial interest(s): T. N. and Y. Y. are employed by Kaneka Corporation. The remaining authors declare no competing financial interests.

ACKNOWLEDGMENTS

We thank Miho Matsumoto-Kitano and Hiromi Oiwa for technical assistance. This work was supported by the Project Focused on Developing Key Technology for Discovering and Manufacturing Drugs for Next-Generation Treatment and Diagnosis from the Japan Agency for Medical Research and Development (AMED), Japan (grant nos. JP20ae0101055, JP20ae0101060, and JP21ae0121002). This work also was partially supported by the Commission for Development of Artificial Gene Synthesis Technology for Creating Innovative Biomaterials from the Ministry of Economy, Trade and Industry (METI), and the JST-Mirai Program (JPMJMI17EJ) and the CREST program (JPMJCR21N2) from the Japan Science and Technology Agency (JST).

ABBREVIATIONS

AOX1, alcohol oxidase 1; cfu, colony-forming units; Dnl4p, DNA ligase IV; *EGFP*, enhanced green fluorescent protein; $G418^R$, G418 resistance marker; *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase; GFP, green fluorescent protein; HR, homologous recombination; NHEJ, non-homologous end joining; ORF, open reading frame; $P_{CCA36283}$, *CCA36283* promoter; P_{GAPDH}, *GAPDH* promoter; P_{TDH1}, *TDH1* promoter; P_{TEF1}, *TEF1* promoter; P_{TEF2}, *TEF2* promoter; T_{38473}, 3′-untranslated (terminator) region of the *CCA38473* gene; T_{AOX1}, *AOX1* terminator; T_{EFM1}, *EFM1* terminator; T_{PRC1}, *PRC1* terminator; T_{RPL3}, *RPL3* terminator

REFERENCES

(1) Gasser, B.; Prielhofer, R.; Marx, H.; Maurer, M.; Nocon, J.; Steiger, M.; Puxbaum, V.; Sauer, M.; Mattanovich, D. *Pichia pastoris*: protein production host and model organism for biomedical research. *Future Microbiol.* **2013**, *8*, 191−208.

(2) Daly, R.; Hearn, M. T. Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.* **2005**, *18*, 119−138.

(3) Cereghino, J. L.; Cregg, J. M. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS Microbiol. Rev.* **2000**, *24*, 45−66.

(4) Ahmad, M.; Hirz, M.; Pichler, H.; Schwab, H. Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 5301−5317.

(5) Zhu, T.; Sun, H.; Wang, M.; Li, Y. Pichia pastoris as a Versatile Cell Factory for the Production of Industrial Enzymes and Chemicals: Current Status and Future Perspectives. *Biotechnol. J.* **2019**, *14*, No. e1800694.

(6) Gasser, B.; Steiger, M. G.; Mattanovich, D. Methanol regulated yeast promoters: production vehicles and toolbox for synthetic biology. *Microb. Cell Fact.* **2015**, *14*, 196.

(7) Vogl, T.; Sturmberger, L.; Kickenweiz, T.; Wasmayer, R.; Schmid, C.; Hatzl, A. M.; Gerstmann, M. A.; Pitzer, J.; Wagner, M.; Thallinger, G. G.; Geier, M.; Glieder, A. A Toolbox of Diverse Promoters Related to Methanol Utilization: Functionally Verified Parts for Heterologous Pathway Expression in Pichia pastoris. *ACS Synth. Biol.* **2016**, *5*, 172−186.

(8) Fischer, J. E.; Glieder, A. Current advances in engineering tools for Pichia pastoris. *Curr. Opin. Biotechnol.* **2019**, *59*, 175−181.

(9) Tadas, Jakočiūnas; Rajkumar, A. S.; Zhang, J.; Arsovska, D.; Rodriguez, A.; Jendresen, C. B.; Skjødt, M. L.; Nielsen, A. T.; Borodina, I.; Jensen, M. K.; Keasling, J. D. CasEMBLR: Cas9-facilitated multiloci genomic integration of in vivo assembled DNA parts in Saccharomyces cerevisiae. *ACS Synth. Biol.* **2015**, *4*, 1226−1234.

(10) De Schutter, K.; Lin, Y. C.; Tiels, P.; Van Hecke, A.; Glinka, S.; Weber-Lehmann, J.; Rouzé, P.; Van de Peer, Y.; Callewaert, N. Genome sequence of the recombinant protein production host Pichia pastoris. *Nat. Biotechnol.* **2009**, *27*, 561−566.

(11) Küberl, A.; Schneider, J.; Thallinger, G. G.; Anderl, I.; Wibberg, D.; Hajek, T.; Jaenicke, S.; Brinkrolf, K.; Goesmann, A.; Szczepanowski, R.; Pühler, A.; Schwab, H.; Glieder, A.; Pichler, H. High-quality genome sequence of Pichia pastoris CBS7435. *J. Biotechnol.* **2011**, *154*, 312−320.

(12) Sturmberger, L.; Chappell, T.; Geier, M.; Krainer, F.; Day, K. J.; Vide, U.; Trstenjak, S.; Schiefer, A.; Richardson, T.; Soriaga, L.;

Darnhofer, B.; Birner-Gruenberger, R.; Glick, B. S.; Tolstorukov, I.; Cregg, J.; Madden, K.; Glieder, A. Refined Pichia pastoris reference genome sequence. *J. Biotechnol.* **2016**, *235*, 121−131.
(13) Nakamura, Y.; Nishi, T.; Noguchi, R.; Ito, Y.; Watanabe, T.; Nishiyama, T.; Aikawa, S.; Hasunuma, T.; Ishii, J.; Okubo, Y.; Kondo, A. A Stable, Autonomously Replicating Plasmid Vector Containing Pichia pastoris Centromeric DNA. *Appl. Environ. Microbiol.* **2018**, *84*, e02882−e02817.
(14) Camattari, A.; Goh, A.; Yip, L. Y.; Tan, A. H.; Ng, S. W.; Tran, A.; Liu, G.; Liachko, I.; Dunham, M. J.; Rancati, G. Characterization of a panARS-based episomal vector in the methylotrophic yeast Pichia pastoris for recombinant protein production and synthetic biology applications. *Microb. Cell Fact.* **2016**, *15*, 139.
(15) Cregg, J. M.; Barringer, K. J.; Hessler, A. Y.; Madden, K. R. Pichia pastoris as a host system for transformations. *Mol. Cell. Biol.* **1985**, *5*, 3376−3385.
(16) Liachko, I.; Dunham, M. J. An autonomously replicating sequence for use in a wide range of budding yeasts. *FEMS Yeast Res.* **2014**, *14*, 364−367.
(17) Schwarzhans, J. P.; Luttermann, T.; Wibberg, D.; Winkler, A.; Hübner, W.; Huser, T.; Kalinowski, J.; Friehs, K. A Mitochondrial Autonomously Replicating Sequence from *Pichia pastoris* for Uniform High Level Recombinant Protein Production. *Front. Microbiol.* **2017**, *8*, 780.
(18) Papakonstantinou, T.; Harris, S.; Hearn, M. T. Expression of GFP using Pichia pastoris vectors with zeocin or G-418 sulphate as the primary selectable marker. *Yeast* **2009**, *26*, 311−321.
(19) Lin Cereghino, G. P.; Lin Cereghino, J.; Sunga, A. J.; Johnson, M. A.; Lim, M.; Gleeson, M. A.; Cregg, J. M. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of Pichia pastoris. *Gene* **2001**, *263*, 159−169.
(20) Vogl, T.; Gebbie, L.; Palfreyman, R. W.; Speight, R. Effect of Plasmid Design and Type of Integration Event on Recombinant Protein Expression in Pichia pastoris. *Appl. Environ. Microbiol.* **2018**, *84*, e02712−e02717.
(21) Invitrogen. (2014) *Pichia expression kit user guide*. Publication MAN0000012. Life Technologies, Carlsbad, CA.
(22) Prielhofer, R.; Barrero, J. J.; Steuer, S.; Gassler, T.; Zahrl, R.; Baumann, K.; Sauer, M.; Mattanovich, D.; Gasser, B.; Marx, H. GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast Pichia pastoris. *BMC Syst. Biol.* **2017**, *11*, 123.
(23) Ma, H.; Kunes, S.; Schatz, P. J.; Botstein, D. Plasmid construction by homologous recombination in yeast. *Gene* **1987**, *58*, 201−216.
(24) Gibson, D. G.; Benders, G. A.; Andrews-Pfannkoch, C.; Denisova, E. A.; Baden-Tillson, H.; Zaveri, J.; Stockwell, T. B.; Brownley, A.; Thomas, D. W.; Algire, M. A.; Merryman, C.; Young, L.; Noskov, V. N.; Glass, J. I.; Venter, J. C.; Hutchison, C. A., 3rd; Smith, H. O. Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science* **2008**, *319*, 1215−1220.
(25) Gibson, D. G.; Benders, G. A.; Axelrod, K. C.; Zaveri, J.; Algire, M. A.; Moodie, M.; Montague, M. G.; Venter, J. C.; Smith, H. O.; Hutchison, C. A., 3rd One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20404−20409.
(26) Kuijpers, N. G.; Chroumpi, S.; Vos, T.; Solis-Escalante, D.; Bosman, L.; Pronk, J. T.; Daran, J. M.; Daran-Lapujade, P. One-step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in Saccharomyces cerevisiae. *FEMS Yeast Res.* **2013**, *13*, 769−781.
(27) Shah, K. A.; Clark, J. J.; Goods, B. A.; Politano, T. J.; Mozdzierz, N. J.; Zimnisky, R. M.; Leeson, R. L.; Love, J. C.; Love, K. R. Automated pipeline for rapid production and screening of HIV-specific monoclonal antibodies using pichia pastoris. *Biotechnol. Bioeng.* **2015**, *112*, 2624−2629.
(28) Royle, K. E.; Polizzi, K. A streamlined cloning workflow minimising the time-to-strain pipeline for Pichia pastoris. *Sci. Rep.* **2017**, *7*, 15817.
(29) Mizutani, K.; Yoshioka, S.; Mizutani, Y.; Iwata, S.; Mikami, B. High-throughput construction of expression system using yeast Pichia pastoris, and its application to membrane proteins. *Protein Expression Purif.* **2011**, *77*, 1−8.
(30) Mizutani, K. High-throughput plasmid construction using homologous recombination in yeast: its mechanisms and application to protein production for X-ray crystallography. *Biosci., Biotechnol., Biochem.* **2015**, *79*, 1−10.
(31) Corrigan, M. W.; Kerwin-Iosue, C. L.; Kuczmarski, A. S.; Amin, K. B.; Wykoff, D. D. The fate of linear DNA in Saccharomyces cerevisiae and Candida glabrata: the role of homologous and non-homologous end joining. *PLoS One* **2013**, *8*, No. e69628.
(32) Schwarzhans, J. P.; Wibberg, D.; Winkler, A.; Luttermann, T.; Kalinowski, J.; Friehs, K. Non-canonical integration events in Pichia pastoris encountered during standard transformation analysed with genome sequencing. *Sci. Rep.* **2016**, *6*, 38952.
(33) Näätsaari, L.; Mistlberger, B.; Ruth, C.; Hajek, T.; Hartner, F. S.; Glieder, A. Deletion of the Pichia pastoris KU70 homologue facilitates platform strain generation for gene expression and synthetic biology. *PLoS One* **2012**, *7*, No. e39720.
(34) Ito, Y.; Watanabe, T.; Aikawa, S.; Nishi, T.; Nishiyama, T.; Nakamura, Y.; Hasunuma, T.; Okubo, Y.; Ishii, J.; Kondo, A. Deletion of DNA ligase IV homolog confers higher gene targeting efficiency on homologous recombination in Komagataella phaffii. *FEMS Yeast Res.* **2018**, *18*, foy074.
(35) Cai, P.; Duan, X.; Wu, X.; Gao, L.; Ye, M.; Zhou, Y. J. Recombination Machinery Engineering facilitates metabolic engineering of the industrial yeast Pichia pastoris. *Nucleic Acids Res.* **2021**, *49*, 7791−7805.
(36) Shinde, D.; Lai, Y.; Sun, F.; Arnheim, N. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Res.* **2003**, *31*, 974−980.
(37) Nishizaki, T.; Tsuge, K.; Itaya, M.; Doi, N.; Yanagawa, H. Metabolic engineering of carotenoid biosynthesis in Escherichia coli by ordered gene assembly in Bacillus subtilis. *Appl. Environ. Microbiol.* **2007**, *73*, 1355−1361.
(38) Vogl, T.; Ruth, C.; Pitzer, J.; Kickenweiz, T.; Glieder, A. Synthetic Core Promoters for Pichia pastoris. *ACS Synth. Biol.* **2014**, *3*, 188−191.
(39) Cheng, T. H.; Chang, C. R.; Joy, P.; Yablok, S.; Gartenberg, M. R. Controlling gene expression in yeast by inducible site-specific recombination. *Nucleic Acids Res.* **2000**, *28*, E108.
(40) Ito, Y.; Terai, G.; Ishigami, M.; Hashiba, N.; Nakamura, Y.; Bamba, T.; Kumokita, R.; Hasunuma, T.; Asai, K.; Ishii, J.; Kondo, A. Exchange of endogenous and heterogeneous yeast terminators in Pichia pastoris to tune mRNA stability and gene expression. *Nucleic Acids Res.* **2020**, *48*, 13000−13012.
(41) Tominaga, M.; Nozaki, K.; Umeno, D.; Ishii, J.; Kondo, A. Robust and flexible platform for directed evolution of yeast genetic switches. *Nat. Commun.* **2021**, *12*, 1846.

https://doi.org/10.1021/acssynbio.1c00302
ACS Synth. Biol. 2022, 11, 644−654