Synthetic and Systems Biotechnology 9 (2024) 759–765

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Screening and characterization of integration sites based on CRISPR-Cpf1 in *Pichia pastoris*

Shupeng Ruan [a,b], Yuxin Yang [a,b], Xinying Zhang [a,b], Guanjuan Luo [a,b], Ying Lin [a,b,**], Shuli Liang [a,b,*]

[a] *Guangdong Key Laboratory of Fermentation and Enzyme Engineering, School of Biology and Biological Engineering, South China University of Technology, Guangzhou, 510006, China*

[b] *Guangdong Research Center of Industrial Enzyme and Green Manufacturing Technology, School of Biology and Biological Engineering, South China University of Technology, Guangzhou, 510006, China*

ARTICLE INFO

Keywords:
P. pastoris
CRISPR-Cpf1
Integration sites
Metabolic engineering

ABSTRACT

Pichia pastoris, a methylotrophic yeast, can utilize methanol as a carbon source and energy source to synthesize high-value chemicals, and is an ideal host for biomanufacturing. Constructing the *P. pastoris* cell factory is somewhat impeded due to the absence of genetic tools for manipulating multi-gene biosynthetic pathways. To broaden its application in the field of metabolic engineering, this study identified and screened 15 novel integration sites in *P. pastoris* using CRISPR-Cpf1 genome editing technology, with EGFP serving the reporter protein. These integration sites have integration efficiencies of 10–100 % and varying expression strengths, which allow for selection based on the expression levels of genes as needed. Additionally, these integrated sites are applied in the heterologous biosynthesis of *P. pastoris*, such as the astaxanthin biosynthetic pathway and the carbon dioxide fixation pathway of the Calvin-Benson-Bassham (CBB) cycle. During the three-site integration process, the 8 genes of the CBB cycle were integrated into the genome of *P. pastoris*. This indicates the potential of these integration sites for integrating large fragments and suggests their successful application in metabolic engineering of *P. pastoris*. This may lead to improved efficiency of genetic engineering in *P. pastoris*.

1. Introduction

P. pastoris is widely recognized as one of the most significant protein expression systems and is extensively utilized for the production of proteins and enzymes [1,2]. *P. pastoris* protein expression offers several key advantages, primarily due to its high secretion capacity [3], and *P. pastoris* exhibits a robust methanol-induced promoter and exceptional posttranslational modification capabilities [4–6]. *P. pastoris* is generally recognized as safe (GRAS) by the Food and Drug Administration, making it suitable for a wide range of applications in the pharmaceutical and food industries. Consequently, there is a growing interest in utilizing *P. pastoris* as a microbial cell factory for the production of both primary and secondary metabolites [7].

Episomal expression vectors and chromosomal integration serve as viable alternatives for expressing heterologous genes. Autonomously replicating sequences (ARSs) have found widespread use in synthetic biology. The most commonly used ARSs in *P. pastoris* are PARS1, mitoARS, and panARS [8]. However, the stability of ARS1-harboring episomal expression vectors in *P. pastoris* was not high unless antibiotic selection pressure was applied [9]. Chromosome integration is the primary method for expressing heterologous genes in *P. pastoris*, as opposed to using episomal expression vectors [10–12]. However, *P. pastoris*, a nonconventional yeast, suffers from limited screening markers and integration sites. Despite the establishment of a recyclable plasmid system for selective screening markers in *P. pastoris* by Li et al. [13], the transition from yeast electroporation to resistance knockout requires approximately 2 weeks, making this a time-consuming process. Furthermore, the scarcity of commonly utilized integration sites in *P. pastoris* poses a significant limitation in the multigenic synthesis pathway. With the introduction of clustered regularly spaced short

Peer review under responsibility of KeAi Communications Co., Ltd.

* Corresponding author.

** Corresponding author. Guangdong Key Laboratory of Fermentation and Enzyme Engineering, School of Biology and Biological Engineering, South China University of Technology, Guangzhou, 510006, China.

E-mail addresses: feylin@scut.edu.cn (Y. Lin), shuli@scut.edu.cn (S. Liang).

https://doi.org/10.1016/j.synbio.2024.06.002

Received 4 April 2024; Received in revised form 11 June 2024; Accepted 11 June 2024

Available online 18 June 2024

palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems in yeast [14], the challenge of limited integration sites in nonconventional yeasts such as *P. pastoris* has been progressively addressed. Zhou et al. pioneered the utilization of CRISPR-Cas9 technology to discover 46 potential genome integration sites in *P. pastoris*. These authors employed neutral sites and promoter libraries to accomplish dual-factor regulation of gene expression and metabolic pathways [15]. Lian et al. developed a CRISPR-based synthetic biology toolkit in *P. pastoris*, achieving integration efficiencies of approximately 100 %, 93 %, and 75 % at single, dual, and triple loci [16], respectively. While CRISPR-Cas9 technology has enabled the identification of certain integration sites in *P. pastoris*, it should be noted that the PAM sequence recognized by Cas9 is 5′ NGG. Moreover, there is a greater frequency of target identification for sites with a greater number of GC base pairs, whereas fewer targets have been identified for sites with a greater number of AT base pairs [17,18]. CRISPR-Cpf1 recognizes a PAM sequence of 5' TTTV [19,20]. Consequently, we attempted to employ CRISPR-Cpf1 technology to identify integration sites that had not been detected using CRISPR-Cas9 technology. Moreover, the application of CRISPR-Cpf1 technology facilitates the exploration of potential sites in *P. pastoris*.

In the present study, we utilized the CRISPR-Cpf1 tool to identify potential neutral sites in *P. pastoris* based on its genome sequence. Green fluorescent protein (*EGFP*) was used as a reporter gene for screening purposes [21]. We additionally assessed the growth performance of strains with integrated genomes under different carbon sources, namely glucose or methanol, as the only carbon source. Finally, these integration sites have proven successful in the biosynthesis of the astaxanthin pathway and the construction of the CBB pathway in *P. pastoris*, highlighting their potential for metabolic production in *P. pastoris*.

2. Materials and methods

2.1. Strains, media, and reagents

E. coli Top10 F′ was employed as the host for plasmid construction. *P. pastoris* GS115 was utilized for gene editing. *E. coli* Top10 F' and the yeast were cultured in LBL media (1 % peptone, 0.5 % yeast extract, and 0.5 % NaCl) at 37 °C and YPD media (1 % yeast extract, 2 % peptone, and 2 % glucose) at 30 °C, respectively. The antibiotics used were as follows: Zeocin (25 mg/L) was utilized for the selection and maintenance of bacterial plasmids in *E. coli*, and 100 mg/L Zeocin was added to YPD media. MD plates (1.34 % YNB (without amino acids), 4×10^{-5} % biotin, 2 % glucose, and 2 % agar) were used for selecting histidine auxotrophs. DNA fragments were amplified using KOD DNA polymerase (TOYOBO, Japan) and Taq DNA polymerase (Thermo Scientific, Waltham, MA). Plasmids were assembled using the NEB Builder HiFi DNA Assembly Master Mix kit (New England BioLabs, Boston, MA). The constructed plasmids were subjected to DNA sequencing for verification (Sangon Biotech, Shanghai, China). Yeast genomic DNA was extracted using a TIAN-amp Yeast DNA kit (TIANGEN), while plasmid DNA was extracted using a TIAN-prep Rapid Mini Plasmid kit (TIANGEN).

2.2. Plasmid construction and strain construction

The plasmids constructed in this study are detailed in Table S1 of the Supporting Information. Oligonucleotides from TSINGKE Biological Technology (Hangzhou, China) were utilized for gRNA construction, DNA amplification, plasmid assembly, and diagnostic PCR verification. Details are provided in Table S2 of the Supporting Information. Plasmids were constructed through restriction digestion/ligation or Gibson assembly methods. The CrRNA plasmid was constructed using the ADE2-crRNA plasmid preserved in the laboratory as a template to generate fragments via PCR. After *Dpn*I digestion, the gene was transformed into Top10 F' and the correct plasmid was obtained through sequencing. The DNA donor was constructed by amplifying the upstream and downstream homologous arms from the yeast genome as the template, followed by overlapping them with the expression box of the target gene to obtain the donor. The PGAP-Cpf1 plasmid was linearized using *Eam1105*I and then electroporated into *P. pastoris* GS115. In GS115-Cpf1, the Cre-loxp technique was used to divide the Ku70 gene and construct the GS115-Cpf1-ΔKu70 strains. The *DAS* gene was deleted from the genome of the GS115 strain using the homologous recombination method. The upper and lower 1000 bp homologous arms of the DAS1 and DAS2 genes were constructed with the His gene using overlap PCR. Positive transformants were then selected in MD medium. Subsequently, the His gene was knocked out using the Cre-loxp method. Typically, 600 ng of crRNA plasmids and 2 μg of linear DNA fragments (DNA donors for integration) were used for electroporation. For the construction of astaxanthin producing strains, we generated the donor fragments P1NS27-*AaCrtZ* and P4NS14*-*HpCrtW*. Next, Cpf1-P1NS27-sgRNA and Cpf1-P4NS14*-sgRNA were introduced into the carotene producing strain *GS115-TmCrtE-SpCrtYB-XdCrtI*. A comprehensive list of all the recombinant strains constructed in this study is provided in Table S3 of the Supporting Information.

2.3. Measurement of relative fluorescence units

The orifice plate culture method was performed following the procedure described by Liao et al. [12], using a 48-well plate. The specific methods were as follows: Strains from the selected plate were pre cultivated in 2 mL 48 deep well plates containing 800 μL of BMGY media (2 % peptone, 1 % yeast extract, 1.34 % YNB, 100 mM potassium phosphate (pH 6.0), 4×10^{-5}% biotin, and 2 % glycerol) for 48 h at 30 °C and 800 rpm on a microtiter plate shaker. Subsequently, 10 μL of the preculture was transferred to another 48-deep-well plate containing 900 μL of BMMY media (2 % peptone, 1 % yeast extract, 1.34 % YNB, 100 mM potassium phosphate (pH 6.0), 4×10^{-5}% biotin, and 1 % methanol) and incubated for 24 h at 30 °C and 900 rpm. After fermentation, 200 μL of the cell fermentation solution was collected and centrifuged to remove the supernatant. The cells were washed three times with 7.4 PBS (8.005 g/L NaCl, 0.2 g/L KCl, 1.42 g/L Na_2HPO_4, 0.245 g/L KH_2PO_4). The EGFP fluorescence intensities were measured at 488–520 nm using a microcoder, and the readings were normalized to the cell density (OD_{600}).

2.4. Measurement of pectate lyase

We have selected the pectin lyase gene *pel-s* (AB428424) from *Bacillus* species and cultivated it using the BMMY fermentation broth system, with methanol supplemented every 24 h. An appropriate volume of fermentation liquid was centrifuged at 10,000 rpm for 5 min. The supernatant was then adequately diluted and 10 μL was added to a 2 ml 50 mM pH 10.0 Gly-NaOH reaction system (containing 0.5 % PGA and 0.5 mM $CaCl_2$) for a 10-min reaction at 80 °C. The reaction was stopped by adding 3 mL of 0.03 M H_3PO_4 solution, and the absorbance at 235 nm was measured. An inactivated enzyme solution was utilized as the blank control group. One unit (U) of pectin lyase activity was defined as the enzyme quantity needed to generate unsaturated oligogalacturonide equivalent to 1 μmol of unsaturated diglucuronide per minute, with a molecular extinction coefficient of 4600 M^{-1} cm^{-1} at 235 nm.

2.5. Analysis and Extraction of products

Astaxanthin was extracted from the fermentation products using a hexane-acetone mixture. The analysis of astaxanthin was conducted using an Agilent 1220 series HPLC system (Agilent Technologies, Germany) equipped with a Sapphire C18 column (4.6 mm × 250 mm), and a UV–vis detector recorded signals at 475 nm. For separation, samples were eluted using an isocratic elution program with a mixture of acetonitrile, isopropanol, and methanol (50:20:30 v/v) at a flow rate of 1 mL/min and maintained at 40 °C.

3. Results

3.1. Screening and characterization of intergenic sites

In our previous studies, we were the first to employ Cpf1 for gene knockout and integration in *P. pastoris* [22]. Currently, the screening and identification of integration sites in *P. pastoris* genome using Cpf1 is carried out. Fig. 1 illustrates the principle of selecting integration sites using Cpf1. Integration sites should not only ensure stable expression of heterologous genes but also minimize their impact on cell growth [23]. Therefore, the regions between two genes are generally selected as candidate integration sites [16]. The genome sequence of *P. pastoris* was used as a reference to screen integration sites. Fig. 1A illustrates the division of intergenic sites into three categories based on the transcription directions of adjacent open reading frames (ORFs): promoter-to-promoter localization (PP), promoter-terminator localization (PT or TP), and terminator-to-terminator localization (TT). Typically, the lengths of promoters and terminators are approximately 750 bp and 500 bp [15], respectively. For spacer sites of the PP type, candidate integration sites with sequence lengths greater than 2100 bp were selected, while screening sites for PAM sites were set at 600 bp. Similarly, for the PT/TP types, the interval region needs to have a sequence length greater than 1850 bp, while for the TT types, the interval region needs to have a sequence length greater than 1600 bp. Based on these criteria, a total of 102 candidate integration sites were identified across the four chromosomes of *P. pastoris* (Fig. 1B). From this pool, based on their lengths, 15 sites were selected from the 102 candidate sites for preliminary validation among three different types of intergenic regions.

Yeast canonically employ two basic mechanisms, namely, homologous recombination (HR) and nonhomologous end joining (NHEJ), to repair DNA double-strand breaks (DSBs) [24]. Regrettably, the efficiency of homologous recombination (HR) in *P. pastoris* has been found to be the limiting factor in genome editing [25]. To attenuate NHEJ and enhance the HR process, the ku70 factor was knocked out in *P. pastoris* [26]. Consequently, the ku70 gene was knocked out in the yeast strain GS115-Cpf1, which was subsequently used as the host strain. The optimized platform achieved integrating efficiencies ranging from 10 % to 100 % when targeting each site with the *eGFP* gene (Table 1). The integration efficiency of the P1NS27 and P4NS14* sites reached 100 %. Furthermore, the genomic locations of some PAM sequences used in these 15 loci are close to the neutral loci documented by Zhou et al. and Lian et al. For instance, the PAM sequence of the P1NS27 site differs with a difference of only about 20bp from the Int1 site (integration efficiency close to 100 %) documented by Lian et al. [16], and it differs with a difference of nearly 500bp from the PNSI-16 site (integration efficiency 80 %) documented by Zhou et al. [15]. Likewise, the PAM sequence of the P4NS14* site differs with a difference of approximately 250bp from the Int18 site (integration efficiency close to 90 %) documented by Lian et al. [16]. Although Cas9 and Cpf1 recognize different PAM sequences, these findings also suggest that near the genomic locations with high Cas9 cleavage efficiency, Cpf1 cleavage efficiency could be high. The expression of each site varies. When assessing the feasibility of integration sites, it is customary to investigate the level of gene expression for the target genes integrated at each site under various carbon sources [27]. When glucose was used as the carbon source, the green fluorescence intensity of P1NS4, P2NS9, and P4NS11 was significantly greater than that of the other sites (Fig. 2B). Under methanol conditions, the eGFP gene expression at the P2NS9 and P4NS11 sites did not differ significantly from that at the other sites, but P1NS4 remained highly expressed (Fig. 2A). This difference in gene expression can be fine-tuned to achieve a balance in the metabolic pathway.

The integration efficiency at each site was evaluated by yeast colony PCR, with a total of thirty individual colonies randomly selected for colony PCR analysis. The integration efficiency is determined by the ratio of positive clones to the total number of selected clones.

3.2. The comparison of expression intensity between the integration site and the commonly used His site

Out of the 15 sites, P1NS4, P2NS9, and P4NS11 exhibit the highest expression intensity under glucose conditions. Therefore, these three sites were selected for comparison with the commonly used His site in *P. pastoris*. The EGFP gene is expressed in cells as a heterologous gene. The growth curve of the PNS strains under glucose conditions (Fig. 3A) was similar to that of GS115-PGAP-EGFP, suggesting that the integration of the EGFP gene into P1NS4, P2NS9, and P4NS11 did not impact yeast growth. Furthermore, at the flask level, the P1NS4 site maintained a

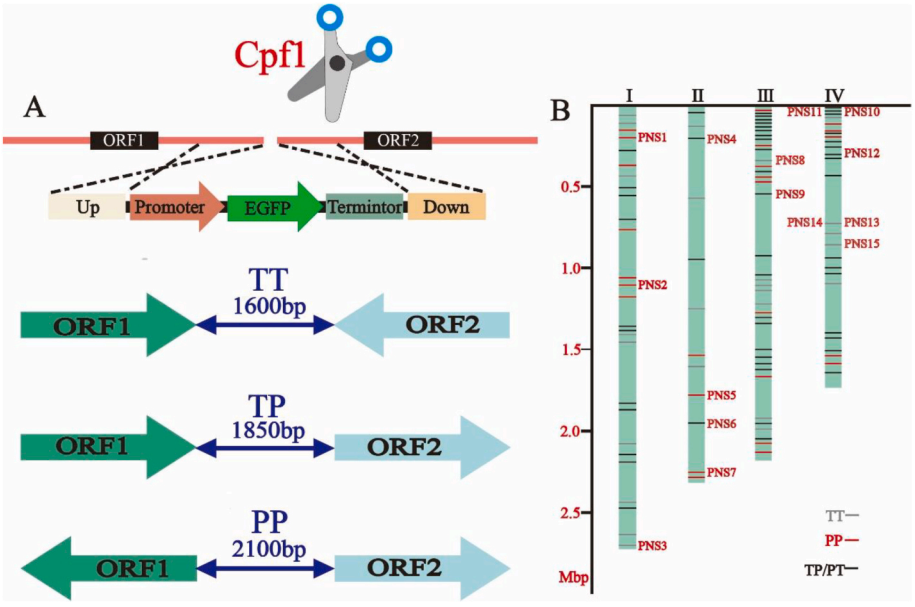

Fig. 1. Screening criteria for candidate integration sites in the *P. pastoris* genome
(A) Protocol for integrating the EGFP gene into candidate integration sites via the CRISPR/Cpf1 system (B) Location of candidate integration sites in the *P. pastoris* genome.

Table 1
Integration efficiency of the EGFP expression box for 15 integration sites.

Candidate neutral site	Flanking gene direction	Chromosome	Position	Guide RNA sequence	Integration efficiency
P1NS4	PP	I	62022~64773	CAAACAATGTTCTGCCTCGTATA	10 %
P1NS13	PP	I	1266752~1268884	TATTGCAAATAATGAAGCTTATC	10 %
P1NS27	TT	I	2788420~2790871	CCCCCAGTAAATACTTCAGATAT	100 %
P2NS3	PT	II	25298~28006	ACATTGACGAGTATGGAAGATTA	20 %
P2NS9	PP	II	1864365~1866886	TCATAGGTGTTGATTCTATAGCG	10 %
P2NS10	PT	II	1918881~1922536	TCTCGGATTGTTCGAACTTGCAA	10 %
P2NS12	PP	II	2378917~2383164	TGGTAGCCACGATACTATCGGCC	20 %
P3NS12	TT	III	308962~311262	AAACGAAGCCAAAGGGCAAATAG	30 %
P3NS17	TP	III	585927~588983	AACGTAGTCACTGGAAGTTGGAC	20 %
P4NS1	TP	IV	6505~9831	CATGGCAAGTTGATTTTAACTCC	10 %
P4NS1*	TP	IV	6505~9831	ATTTACGTTCTCAGTAATGCGAC	20 %
P4NS11	TP	IV	334809~340774	AGGATGACTGCGTGAGATCTGTA	10 %
P4NS14	TT	IV	732400~734110	AACACCCATAAACACTTGTCATT	70 %
P4NS14*	TT	IV	732400~734110	TCTCTCTAACATTCACCTGATCT	100 %
P4NS16	TT	IV	908939~911936	TGAGACGATGATTTCGAAAGGTT	10 %

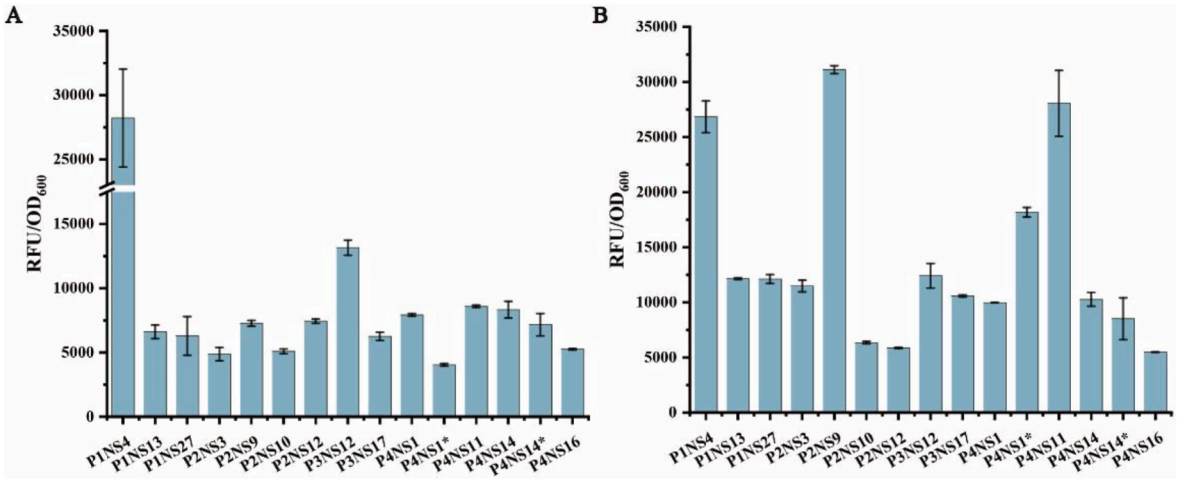

Fig. 2. Fluorescence intensity of 15 PNS strains in a 48-well plate
(A) methanol as a carbon source (B) glucose as a carbon source.

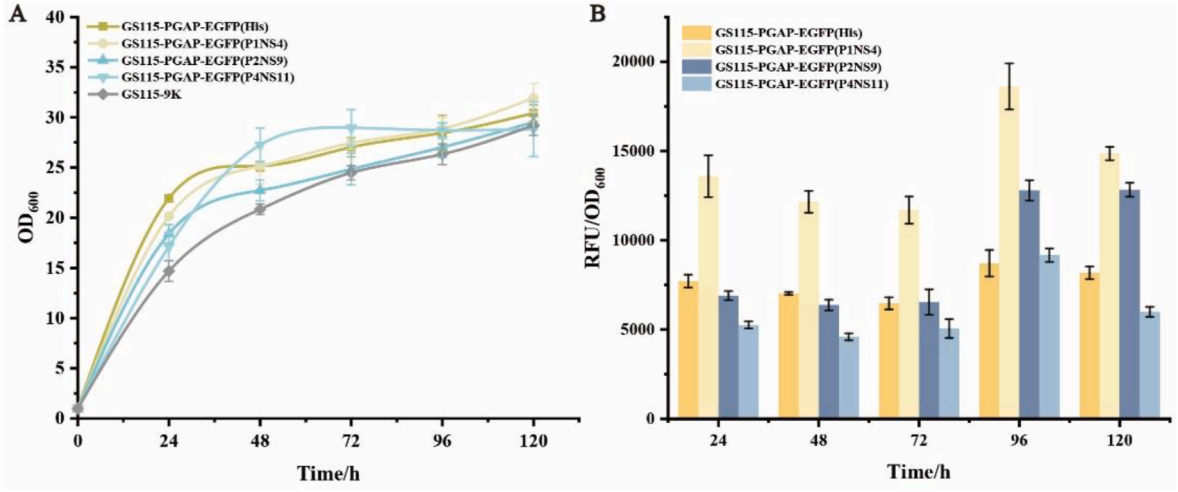

Fig. 3. Fluorescence intensity of the P1NS4,P2NS9 and P4NS11 strains in glucose medium
(A) Growth curve (B) Relative fluorescence intensity.

consistently high expression level, while P2NS9 did not surpass the expression level of the control strain GS115-PGAP-EGFP until the fourth day of fermentation. However, the expression intensity of P4NS11 was similar to that of the control strain (Fig. 3B). Because EGFP demonstrates the highest intracellular expression efficiency under methanol conditions at the P1NS4 site, an attempt was made to utilize the P1NS4 site for secreting the pectinase protein, followed by testing the enzyme activity levels at P1NS4 site. In contrast to the results of intracellular expression,

P1NS4 strains with integrated pectinase at commonly used His sites were used as positive controls for pectinase secretion. It is evident that the growth environment of the strain producing pectinase at P1NS4 site is basically the same as that of the strain producing pectinase at the His site (Fig. 4A). However, the enzymatic activity of the strain producing pectinase at P1NS4 site is only 60 % of that of the strain producing pectinase at the His site (Fig. 4B). In conclusion, the intracellular protein expression effect of the P1NS4 site is superior to that of His, but in regard to the secreted protein, the P1NS4 site only achieves 60 % of the integration seen with His. This indicates that these sites may be more suitable than other sites for cells to construct pathways for certain metabolites.

3.3. Metabolite production by use of neutral sites

We used the *β-carotene* strain as a host and integrated the *AaCrtZ* and *HpCrtW* genes into the neutral sites of P1NS27 and P4NS14*, respectively, to obtain the intermediate products zeaxanthin and kerataxanthin and the end product astaxanthin (Fig. 5A). Finally, 13.29 mg/L astaxanthin was produced by the engineered yeast in the media supplemented with methanol as the carbon source (Fig. 5B). The successful construction of the astaxanthin synthesis pathway proved the feasibility of the neutral sites screened in this study in the metabolic engineering of *P. pastoris* and expanded the genetic manipulation methods for synthetic biology research on methyl-trophic yeasts [28,29]. The β-carotene strain was utilized as the host, and the *AaCrtZ* and *HpCrtW* genes were integrated into the neutral sites of P1NS27 and P4NS14*, resulting in the production of the intermediate products zeaxanthin and kerataxanthin, as well as the end product astaxanthin (Fig. 5D). Ultimately, the engineered yeast produced 13.29 mg/L astaxanthin in the medium supplemented with methanol as the carbon source. The successful establishment of the astaxanthin synthesis pathway demonstrated the viability of the neutral sites identified in this study for metabolic engineering in *P. pastoris*, and broadened the scope of genetic manipulation tools for synthetic biology investigations of methylotrophic yeasts.

3.4. CO_2 fixation pathway integration by multisite integration

The increasing concentration of atmospheric carbon dioxide resulting from human activities surpassed the symbolically significant threshold of 400 ppm in 2015, indicating that the world is experiencing a significant climate change crisis [30,31]. Microbial CO_2 fixation offers a promising approach for achieving “carbon neutrality” by employing microbial cell factories as catalysts for CO_2 conversion. This method has the advantages of promoting environmental sustainability and minimizing energy consumption [32].

The current pathway for carbon dioxide fixation consists of multiple enzymes that facilitate the conversion of carbon dioxide into biomass within the cell. However, relying on a single integration site is insufficient for the proper integration of these enzymes in *P. pastoris*. Integrating multiple genes at the same location may lead to the loss or detachment of the integrated gene. Thus, we aimed to fix carbon dioxide in *P. pastoris* by integrating the Calvin–Benson–Bassham (CBB) cycle [33] at three specific sites: P1NS4, P1NS27, and P4NS14*. *RuBisCO*, *GroEL*, and *GroES* were integrated at P1NS4, while P1NS27 integrated *TDH3*, *PRK*, and *PGK1*. P4NS14* integrates *TKL1* and *TPI1* to establish a comprehensive CBB pathway in *P. pastoris*. Simultaneously, we disrupted the *DAS1* and *DAS2* genes in the methanol assimilation pathway of *P. pastoris*, limiting the strain’s ability to utilize methanol solely for the assimilation pathway, resulting in the production of NADH and the conversion of carbon dioxide into biomass via the CBB pathway (Fig. 6A). Ultimately, the autotrophic *P. pastoris* strain was subjected to a 240-h fermentation process in YNM medium, resulting in an OD_{600} increase of approximately 8. However, the bacteria lacking the CBB pathway did not exhibit growth until the methanol assimilation pathway was disrupted (Fig. 6B). Thus, this further demonstrated the correct integration and expression of the 8 genes involved in the CBB pathway at the designated sites, efficiently converting inorganic carbon dioxide into organic carbon, thereby facilitating the growth of *P. pastoris*.

4. Discussion

Despite the widespread use of *P. pastoris* for recombinant protein expression, there is a growing interest in transforming this methylotrophic yeast into a cellular factory for the production of high-value products. The availability of convenient genetic tools is essential for enabling extensive metabolic rewiring and the development of robust microbial cell factories. The genetic manipulation of *P. pastoris* has been facilitated through the utilization of CRISPR-Cas9 genome editing tools [34], and the limited availability of genomic integration sites can impede the metabolic recombination of biosynthetic chemicals. Currently, there are also many reports utilizing Cas9 technology to screen suitable integration sites in *P. pastoris*[15 16]. The PAM sequence of Cas9 is 5′ NGG, resulting in more recognition targets with sequences rich in GC base pairs and fewer targets with genome sequences rich in AT base pairs. In contrast, Cpf1 recognizes a 5’ TTTV PAM sequence, effectively compensating for this limitation.

By analyzing the characteristics of interval sites using CRISPR-Cpf1

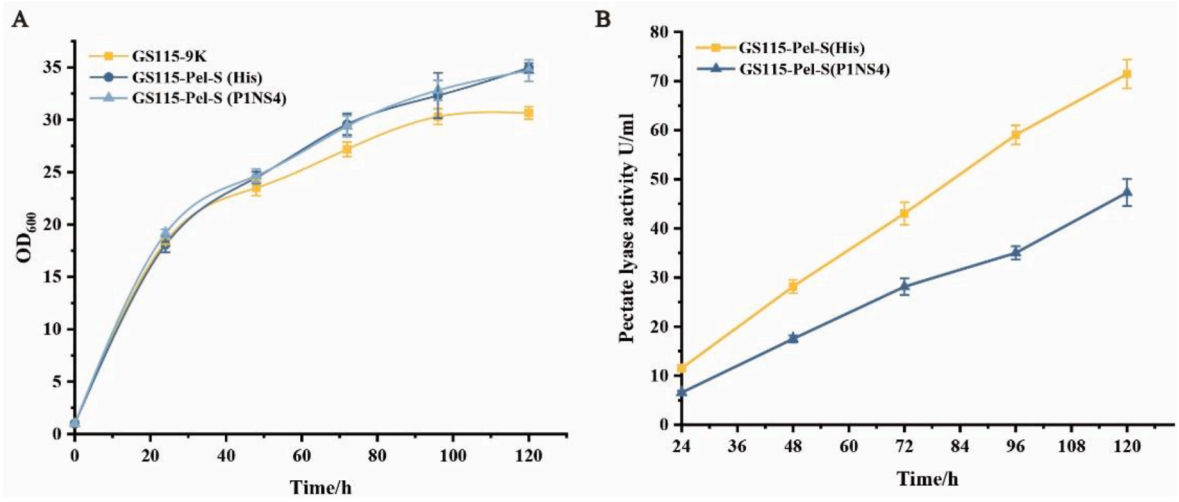

Fig. 4. The GS115-Pel-S (P1NS4) strain was fermented with methanol as a carbon source (A) Growth curve (B) Pectate lyase activity.

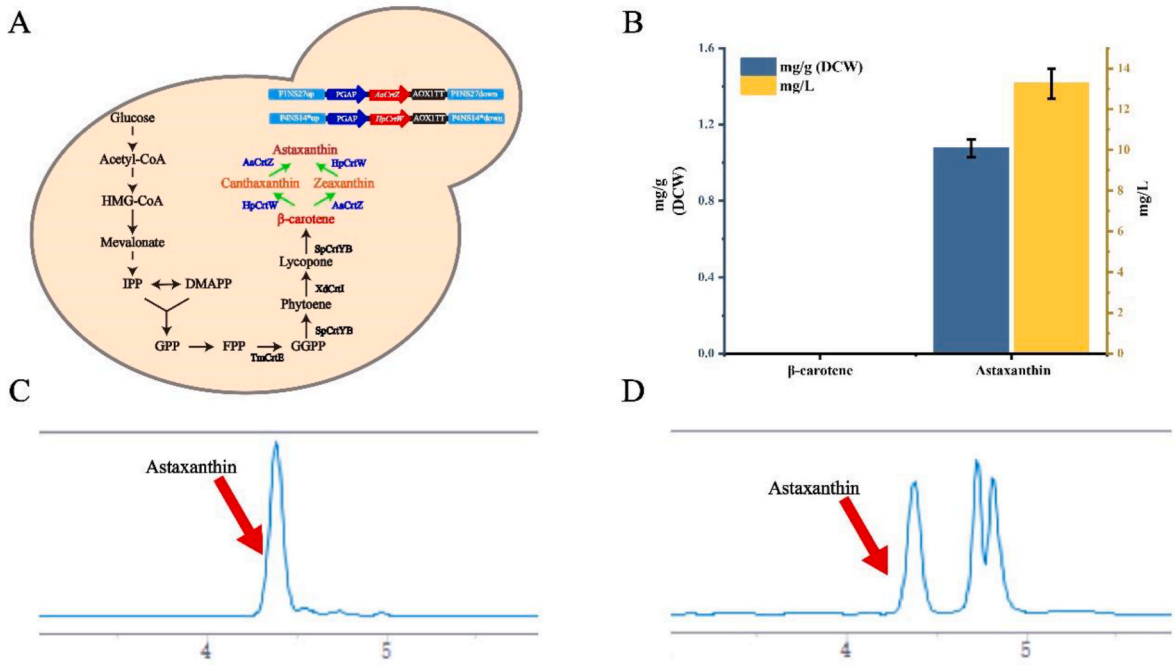

Fig. 5. An astaxanthin biosynthesis pathway was constructed using integration sites
(A) Designed astaxanthin biosynthetic pathway. (B) Astaxanthin production rate and concentration produced by *β-carotene*-engineered yeast and *astaxanthin*-engineered yeast. (C) Liquid chromatogram of the astaxanthin standard. The x-axis is measured in minutes. (D) Liquid chromatography of astaxanthin production by the engineered *P. pastoris* strain. The x-axis is measured in minutes.

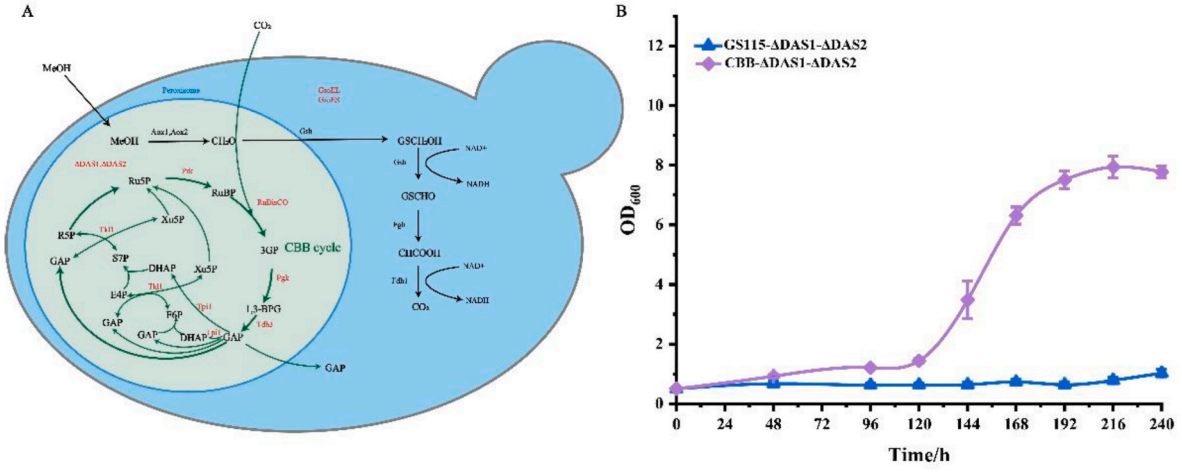

Fig. 6. Construction process for the yeast strain CBB-ΔDAS1-ΔDAS2
(A) Metabolic flow chart (B) Growth curve of the yeast strain CBB-ΔDAS1-ΔDAS2 in flasks.

technology, 15 neutral sites were identified. These 15 sites provide sufficient resources for studying metabolic engineering and synthetic biology in *P. pastoris*. In summary, these neutral sites, which do not hinder cell growth, can be utilized for gene integration in *P. pastoris*. Analysis of the expression characteristics of the 15 sites revealed that P1NS4 exhibited high expression levels when glucose and methanol were used as carbon sources. Furthermore, under glucose conditions, the relative fluorescence intensity of the integrated *EGFP* was greater at the P1NS4, P2NS9, and P4NS11 sites than at the other sites. Such gene expression variation can be employed to finely regulate the expression of related genes in balanced metabolic pathways. Integration of the astaxanthin synthesis pathway genes *CrtZ* and *CrtW* into P1NS27 and P4NS14* resulted in an engineered strain producing 13.29 mg/L of astaxanthin, demonstrating the applicability of the abovementioned neutral site in *P. pastoris* metabolic engineering. The integration of three loci was performed in *P. pastoris* for the construction of the CBB pathway, resulting in the transition of *P. pastoris* from heterotrophic yeast to autotrophic yeast. Additionally, the potential of integrating multiple gene fragments at a single site was explored.

In summary, we conducted a systematic characterization of the spacer region of *P. pastoris* using CRISPR-Cpf1, resulting in the identification of 15 integration sites. These neutral sites will contribute to the development of stable microbial cell factories for biofuturing.

CRediT authorship contribution statement

Shupeng Ruan: designed the research, performed the experiments, analyzed the data and wrote the manuscript, discussed the results and

contributed to theimprovement of the manuscript. **Yuxin Yang:** performed the experiments. **Xinying Zhang:** performed the experiments. **Guanjuan Luo:** performed the experiments. **Ying Lin:** provided guidance and knowledge, discussed the results and contributed to theimprovement of the manuscript, All authors read and approved the final manuscript. **Shuli Liang:** designed the research, provided guidance and knowledge, discussed the results and contributed to theimprovement of the manuscript.

Declaration of competing interest

The authors declare no competing financial interest.

Acknowledgements

This work was supported by the National Key Research and Development Program (2021YFC2104000), the National Natural Science Foundation of China (32272276), and the Fundamental Research Funds for the Central Universities.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2024.06.002.

References

[1] Zahrl RJ, Peña DA, Mattanovich D, Gasser B. Systems biotechnology for protein production in *Pichia pastoris*. FEMS Yeast Res 2017 Nov 1;17(7). https://doi.org/10.1093/femsyr/fox068.
[2] Barone GD, Emmerstorfer-Augustin A, Biundo A, Pisano I, Coccetti P, Mapelli V, Camattari A. Industrial production of proteins with *Pichia pastoris*-Komagataella phaffii. Biomolecules 2023 Feb 26;13(3):441. https://doi.org/10.3390/biom13030441.
[3] Thak EJ, Yoo SJ, Moon HY, Kang HA. Yeast synthetic biology for designed cell factories producing secretory recombinant proteins. FEMS Yeast Res 2020 Mar 1;20(2):foaa009. https://doi.org/10.1093/femsyr/foaa009.
[4] Gasser B, Steiger MG, Mattanovich D. Methanol regulated yeast promoters: production vehicles and toolbox for synthetic biology. Microb Cell Fact 2015 Dec 2;14:196. https://doi.org/10.1186/s12934-015-0387-1.
[5] Puxbaum V, Mattanovich D, Gasser B. Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. Appl Microbiol Biotechnol 2015 Apr;99(7):2925–38. https://doi.org/10.1007/s00253-015-6470-z.
[6] Yang Z, Zhang Z. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: a review. Biotechnol Adv 2018 Jan-Feb;36(1):182–95. https://doi.org/10.1016/j.biotechadv.2017.11.002.
[7] Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. Metabolic engineering of *Pichia pastoris*. Metab Eng 2018 Nov;50:2–15. https://doi.org/10.1016/j.ymben.2018.04.017.
[8] Gu Y, Gao J, Cao M, Dong C, Lian J, Huang L, Cai J, Xu Z. Construction of a series of episomal plasmids and their application in the development of an efficient CRISPR/Cas9 system in *Pichia pastoris*. World J Microbiol Biotechnol 2019 May 27;35(6):79. https://doi.org/10.1007/s11274-019-2654-5.
[9] Liachko I, Dunham MJ. An autonomously replicating sequence for use in a wide range of budding yeasts. FEMS Yeast Res 2014 Mar;14(2):364–7. https://doi.org/10.1111/1567-1364.12123.
[10] Mikkelsen MD, Buron LD, Salomonsen B, Olsen CE, Hansen BG, Mortensen UH, Halkier BA. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metab Eng 2012 Mar;14(2):104–11. https://doi.org/10.1016/j.ymben.2012.01.006.
[11] Kong S, Yu W, Gao N, Zhai X, Zhou YJ. Expanding the neutral sites for integrated gene expression in *Saccharomyces cerevisiae*. FEMS Microbiol Lett 2022 Sep 14;369(1):fnac081. https://doi.org/10.1093/femsle/fnac081.
[12] Liao X, Li L, Jameel A, Xing XH, Zhang C. A versatile toolbox for CRISPR-based genome engineering in *Pichia pastoris*. Appl Microbiol Biotechnol 2021 Dec;105(24):9211–8. https://doi.org/10.1007/s00253-021-11688-y.
[13] Li C, Lin Y, Zheng X, Yuan Q, Pang N, Liao X, Huang Y, Zhang X, Liang S. Recycling of a selectable marker with a self-excisable plasmid in *Pichia pastoris*. Sci Rep 2017 Sep 11;7(1):11113. https://doi.org/10.1038/s41598-017-11494-5.
[14] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res 2013 Apr;41(7):4336–43. https://doi.org/10.1093/nar/gkt135. Epub 2013 Mar 4.
[15] Cai P, Duan X, Wu X, Gao L, Ye M, Zhou YJ. Recombination machinery engineering facilitates metabolic engineering of the industrial yeast *Pichia pastoris*. Nucleic Acids Res 2021 Jul 21;49(13):7791–805. https://doi.org/10.1093/nar/gkab535.
[16] Gao J, Xu J, Zuo Y, Ye C, Jiang L, Feng L, Huang L, Xu Z, Lian J. Synthetic biology toolkit for marker-Less integration of Multigene pathways into *Pichia pastoris* via CRISPR/Cas9. ACS Synth Biol 2022 Feb 18;11(2):623–33. https://doi.org/10.1021/acssynbio.1c00307.
[17] Rainha J, Rodrigues JL, Rodrigues LR. CRISPR-Cas9: a Powerful tool to efficiently engineer *Saccharomyces cerevisiae*. Life 2020 Dec 26;11(1):13. https://doi.org/10.3390/life11010013.
[18] Ferreira P, Choupina AB. CRISPR/Cas9 a simple, inexpensive and effective technique for gene editing. Mol Biol Rep 2022 Jul;49(7):7079–86. https://doi.org/10.1007/s11033-022-07442-w.
[19] Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 2015 Oct 22;163(3):759–71. https://doi.org/10.1016/j.cell.2015.09.038.
[20] Swiat MA, Dashko S, den Ridder M, Wijsman M, van der Oost J, Daran JM, Daran-Lapujade P. FnCpf1: a novel and efficient genome editing tool for *Saccharomyces cerevisiae*. Nucleic Acids Res 2017 Dec 1;45(21):12585–98. https://doi.org/10.1093/nar/gkx1007.
[21] De Schutter K, Lin YC, 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 Jun;27(6):561–6. https://doi.org/10.1038/nbt.1544.
[22] Zhang X, Gu S, Zheng X, Peng S, Li Y, Lin Y, Liang S. A novel and efficient genome editing tool Assisted by CRISPR-Cas12a/Cpf1 for *Pichia pastoris*. ACS Synth Biol 2021 Nov 19;10(11):2927–37. https://doi.org/10.1021/acssynbio.1c00172.
[23] Liu T, Gou Y, Zhang B, Gao R, Dong C, Qi M, Jiang L, Ding X, Li C, Lian J. Construction of ajmalicine and sanguinarine de novo biosynthetic pathways using stable integration sites in yeast. Biotechnol Bioeng 2022 May;119(5):1314–26. https://doi.org/10.1002/bit.28040.
[24] Klinner U, Schäfer B. Genetic aspects of targeted insertion mutagenesis in yeasts. FEMS Microbiol Rev 2004 May;28(2):201–23. https://doi.org/10.1016/j.femsre.2003.10.002.
[25] Gassler T, Heistinger L, Mattanovich D, Gasser B, Prielhofer R. CRISPR/Cas9-Mediated Homology-Directed genome editing in *Pichia pastoris*. Methods Mol Biol 2019;1923:211–25. https://doi.org/10.1007/978-1-4939-9024-5_9.
[26] Näätsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A. Deletion of the *Pichia pastoris* KU70 homologue facilitates platform strain generation for gene expression and synthetic biology. PLoS One 2012;7(6):e39720. https://doi.org/10.1371/journal.pone.0039720.
[27] Yu W, Gao J, Zhai X, Zhou YJ. Screening neutral sites for metabolic engineering of methylotrophic yeast *Ogataea polymorpha*. Synth Syst Biotechnol 2021 Mar 31;6(2):63–8. https://doi.org/10.1016/j.synbio.2021.03.001.
[28] Wang DN, Feng J, Yu CX, Zhang XK, Chen J, Wei LJ, Liu Z, Ouyang L, Zhang L, Hua Q, Liu F. Integrated pathway engineering and transcriptome analysis for improved astaxanthin biosynthesis in *Yarrowia lipolytica*. Synth Syst Biotechnol 2022 Aug 18;7(4):1133–41. https://doi.org/10.1016/j.synbio.2022.08.001.
[29] Nair A, Ahirwar A, Singh S, Lodhi R, Lodhi A, Rai A, Jadhav DA, Harish Varjani S, Singh G, Marchand J, Schoefs B, Vinayak V. Astaxanthin as a king of Ketocarotenoids: Structure, synthesis, Accumulation, Bioavailability and Antioxidant Properties. Mar Drugs 2023 Mar 13;21(3):176. https://doi.org/10.3390/md21030176.
[30] Onyeaka H, Ekwebelem OC. A review of recent advances in engineering bacteria for enhanced CO2 capture and utilization. Int J Environ Sci Technol 2023;20(4):4635–48. https://doi.org/10.1007/s13762-022-04303-8.
[31] Kajla S, Kumari R, Nagi GK. Microbial CO2 fixation and biotechnology in reducing industrial CO2 emissions. Arch Microbiol 2022 Jan 21;204(2):149. https://doi.org/10.1007/s00203-021-02677-w.
[32] Guo F, Qiao Y, Xin F, Zhang W, Jiang M. Bioconversion of C1 feedstocks for chemical production using *Pichia pastoris*. Trends Biotechnol 2023 Aug;41(8):1066–79. https://doi.org/10.1016/j.tibtech.2023.03.006.
[33] Gassler T, Sauer M, Gasser B, Egermeier M, Troyer C, Causon T, Hann S, Mattanovich D, Steiger MG. The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO2. Nat Biotechnol 2020 Feb;38(2):210–6. https://doi.org/10.1038/s41587-019-0363-0.
[34] Park J, Kim IJ, Kim SR. Nonconventional yeasts engineered using the CRISPR-Cas system as Emerging microbial cell factories. Fermentation 2022;8:656. https://doi.org/10.3390/fermentation8110656.