Development of synthetic biology tools to engineer *Pichia pastoris* as a chassis for the production of natural products

Jucan Gao [a], Lihong Jiang [a], Jiazhang Lian [a,b,*]

[a] *Key Laboratory of Biomass Chemical Engineering of Ministry of Education, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, 310027, China*
[b] *Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou, 310027, China*

ARTICLE INFO

ABSTRACT

The methylotrophic yeast *Pichia pastoris* (a.k.a. *Komagataella phaffii*) is one of the most commonly used hosts for industrial production of recombinant proteins. As a non-conventional yeast, *P. pastoris* has unique biological characteristics and its expression system has been well developed. With the advances in synthetic biology, more efforts have been devoted to developing *P. pastoris* into a chassis for the production of various high-value compounds, such as natural products. This review begins with the introduction of synthetic biology tools for the engineering of *P. pastoris*, including vectors, promoters, and terminators for heterologous gene expression as well as Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated System (CRISPR/Cas) for genome editing. This review is then followed by examples of the production of value-added natural products in metabolically engineered *P. pastoris* strains. Finally, challenges and outlooks in developing *P. pastoris* as a synthetic biology chassis are prospected.

1. Introduction

Natural products generally refer to secondary metabolites isolated from animals, plants, and microorganisms, which have special physiological activities. Many high-value natural products play an extremely important role in medicine, cosmetics, and industry [1]. However, natural products generally have complex chemical structures and are extracted from sources with long growth periods and extremely low abundance, resulting in short supply and high cost, particularly those with plant origins [2]. With the development of synthetic biology, various microorganisms including *Pichia pastoris* (a.k.a. *Komagataella phaffii*) have been developed as cell factories to produce natural products [3–5]. Compared with bacterial cell factories (i.e. *Escherichia coli*), the ability of post-translational modifications and the presence of inner membrane systems make yeasts including *P. pastoris* preferred hosts to express eukaryotic complex proteins, such as cytochrome P450s (CYPs), which are often involved in the biosynthesis of natural products [6]. When compared with the model yeast *Saccharomyces cerevisiae*, *P. pastoris* has the advantage of strong and tightly regulated promoters for high level expression of recombinant proteins [7]. For example, the

expression level of the target gene can account for more than 30% of the total proteins of *P. pastoris*, which is much higher than that in *S. cerevisiae* [8]. In addition, hyperglycosylation is another concern for expressing eukaryotic proteins in the *S. cerevisiae* expression system [9]. Natural product biosynthetic pathway enzymes (i.e. polyketide synthases and CYPs) are generally found to have relatively low catalytic activities, which should be expressed at high levels to achieve efficient biosynthesis. Therefore, *P. pastoris* is a promising host for large-scale production of natural products, particularly those with eukaryotic origins.

Although generally considered as a non-conventional yeast, genetics, physiology, and cell biology of *P. pastoris* have been studied in-depth [4]. The genomes of *P. pastoris* GS115 and CBS7435 have been sequenced and annotated, and genome-scale metabolic models have been constructed by analyzing the metabolic patterns [10]. Metabolomics studies indicate that the intermediate metabolites of *P. pastoris* and *S. cerevisiae* are very similar, with identical metabolites up to 90% [11]. The exploration of the genetic background of *P. pastoris* has laid a solid foundation for customized modification of *P. pastoris*. Currently, besides therapeutic proteins and enzymes [12], *P. pastoris* has been

engineered to produce various chemicals and value-added compounds, such as D-lactic acid [13], 2,3-butanediol (BDO) [14], 2-phenylethanol [15], isobutanol and isobutyl acetate [16], carotenoids [17], lovastatin [18], and nootkatone [19] (Fig. 1).

In this review, the synthetic biology tools needed for the construction and optimization of natural product biosynthetic pathways in *P. pastoris* are firstly summarized, in particular the heterologous gene expression system and the CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated System) based genome editing system. Subsequently, several examples of the establishment of *P. pastoris* as cell factories for the production of terpenoids, polyketides, and flavonoids are introduced. Finally, future perspectives in the development of novel synthetic biology tools for the assembly and integration of multi-gene biosynthetic pathways and high throughput genome engineering are discussed.

2. Synthetic biology toolkit for *P. pastoris*

2.1. Gene expression vectors

The most common way to introduce exogenous genes into *P. pastoris* is to construct recombinant vectors. The method of plasmid maintenance in yeast is through auxotrophic markers or resistance selection markers (Table 1) [20]. Plasmids can be divided into episomal plasmids and integrative plasmids according to the way existing in the host. Unfortunately, the episomal plasmids suffer from low stability and genome integration is generally preferred for high level expression of heterologous genes in *P. pastoris*. Usually the vector is linearized and integrated into the *P. pastoris* genome in a single copy manner. For example, the vectors pPIC9K-His and pHIL-S1 can be linearized by *Sal*I for the integration into the *HIS4* locus of *P. pastoris* GS115, which are then screened in histidine-dropout medium to obtain single-copy integrated strains [21,22]. In addition to single-copy integration, multi-copy integration is generally demanded for high-level expression of the target proteins. The pPIC series of vectors are commonly used integrative vectors in *P. pastoris* [23–26], which enable the screening of the multi-copy integration strains under high concentration of antibiotics, a mechanism known as post-transformation amplification [27–30]. In addition to the formation of tandem repeats via post-transformation amplification, multi-copy strains can be constructed by integrating into the repetitive sequences of the *P. pastoris* genome, such as the ribosomal DNA (rDNA) sequences [13].

Nevertheless, episomal expression possesses unique advantages for several applications, such as the combinatorial optimization of multi-gene biosynthetic pathways and the development of efficient CRISPR-based genome editing tools [31]. In this case, a set of episomal vectors with various autonomously replicating sequences (ARSs) have been constructed and systematically compared for their transformation efficiency, copy numbers, and reproductive stability (Table 1) [32]. Of a particular note, panARS, a broad host ARS derived from *Kluyveromyces lactis*, was found to enable the highest plasmid stability and chosen for the development of an efficient CRISPR/Cas9 system for *P. pastoris* [33].

2.2. Promoters and terminators

Promoters are considered as the most important synthetic biology elements and have direct impacts on the expression of the transcription units. The selection of appropriate promoters with the desirable strength is essential to construct well-controlled synthetic biology modules and to achieve optimal expression of the target genes. The alcohol oxidase 1 promoter (*pAOX1*) and the glyceraldehyde 3-phosphate dehydrogenase promoter (*pGAP*) are two most commonly used promoters [49]. The *AOX1* promoter is generally considered as the strongest promoter of *P. pastoris*, which is strongly induced by methanol and inhibited by glycerol, ethanol, and glucose. Under the full induction conditions, Aox1p accounted for more than 30% of the total cellular proteins (Table 2) [50]. The *GAP* promoter is a strong constitutive promoter, whose expression strength is relatively stable. The expression level of some foreign proteins under the control of *pGAP* can reach up to the level of g/L [51].

To explore metabolic engineering and synthetic biology applications, a series of promoters with different properties have been characterized and engineered [52]. Based on RNA-seq results, *p0188* (Pyruvate decarboxylase isozyme, locus tag: *PAS_chr3_0188*) was determined to be the strongest constitutive promoter and Aslan et al. found that *p0188* had a strong driving force and the strength could reach up to 2-fold as that of *pGAP* [53]. By combining RNA-seq and mRNA folding free energy, 16 promoter candidates were selected and characterized with glucose, glycerol, or methanol as the sole carbon source, respectively. The *p0547* (peroxidase promoter, locus tag: *PAS_chr1-4_0547*) and *p0472* (mitochondrial alcohol dehydrogenase isozyme III, locus tag: *PAS_chr2-1_0472*) were determined to be the strongest methanol-inducible and constitutive promoters to express α-amylase, respectively [49]. Interestingly, in another separate study, Karaoglan et al. found that the *ADH3* promoter (locus tag: *PAS_chr2-1_0472*) performed even better than the *AOX1* promoter when using methanol as the sole carbon source to express *Aspergillus niger* xylanase [54].

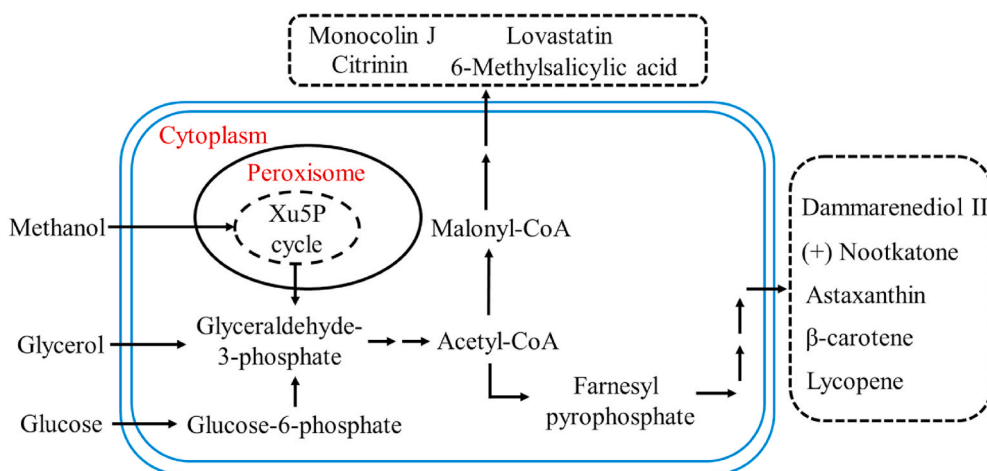In addition to the mining and characterization of endogenous

Fig. 1. *P. pastoris* as a cell factory for the production of natural products. *P. pastoris* converts various carbon sources (i.e. methanol, glycerol, and glucose) into a few central metabolites (i.e. pyruvate and acetyl-CoA), which serve as the precursors for the biosynthesis of a variety of natural products (i.e. terpenoids, polyketides, and flavonoids).

Table 1

Gene expression vectors commonly used in *P. pastoris*.

Plasmid backbone	ARS or integration locus	Copy numbers	Selection markers	Episomal or Integrative	Reference
pHIL-S1[a]	*HIS4*	1.0	*HIS4*	Integrative	[21]
pPIC9K[a,b] or pPIC9K-His[a,b]	*HIS4/AOX1*	1.0 or 4.0–6.0	*HIS4*/G418[r]	Integrative	[23,34–37]
pPIC3.5K[b]	*HIS4/AOX1*	1.0	*HIS4*/G418[r]	Integrative	[24]
pGAPZ[b,c]	*GAP*	1.0 or 21	Zeocin[r]	Integrative	[14,27,38]
pPICZ	*AOX1*	1.0 or 10.0	Zeocin[r]	Integrative	[25,26,28]
pAO815	*HIS4/AOX1*	1.0	*HIS4*	Integrative	[39]
pPink HC or pPink LC	*TRP2*	2.0–4.0	*ADE2*	Integrative	[40,41]
pGLY2664[b]	*TRP2*	2.0–12.0	*ADE2*/Zeocin[r]	Integrative	[42]
pUC19[c]	rDNA	~9.0	Zeocin[r]	Integrative	[13]
pPICZα[a,b]	*AOX1*	~10.0	Zeocin[r]	Integrative	[29]
pMCO[a]	*HIS4*	2.0–16.0	G418[r]	Integrative	[30]
pGHYB	*GAP*	~3.0	Hygromycin[r]	Integrative	[43]
pPIC9K[d]	panARS	~18.0	*HIS4*	Episomal	[32]
pPIC9K[d]	PARS1	~15.0	*HIS4*	Episomal	[32]
pPIC9K[d]	mitoARS	~14.0	*HIS4*	Episomal	[32]
pPIC9K[d]	PpARS2	~4.0	*HIS4*	Episomal	[32]
pPIC9K[d]	ScARS	~4.0	*HIS4*	Episomal	[32]
pSEC-SUMO	panARS	~19.0	Zeocin[r]	Episomal	[33]
pMito	mtDNA	~3.0	*HIS4*	Episomal	[44]
pPICZαBHF[a]	PARS1	~7.0	Zeocin[r]	Episomal	[45]
pBGP1[a]	PARS1	N.A.	Zeocin[r]	Episomal	[46,47]
pPEHα[a]	PARS1	N.A.	*HIS4*	Episomal	[48]

Note.

[a] These plasmids contain a signal peptide for protein secretion.

[b] These *P. pastoris* integrative plasmids have both auxotrophic markers (i.e. *HIS4* and *ADE2*) for single-copy integration and resistance markers (i.e. G418[r] and Zeocin[r]) for multi-copy integration. High-copy strains can be generally constructed by screening on high concentration of antibiotics.

[c] High-copy strains (up to 21 copies) were constructed by integrating the expression cassettes into the ribosomal DNA loci and screening under high Zeocin concentration conditions. Insert the *HIS4* gene into the plasmid to obtain a single copy strain.

[d] With pPIC9K as the backbone, a series of episomal plasmids were constructed by adding different yeast replicons (ARSs).

Table 2

P. pastoris promoters commonly used for the expression of heterologous genes.

Promoter	Type	Relative Strength[a]		Locus tag	Reference
		Methanol	Glucose		
pAOX1	Methanol inducible	~18	/	*PAS_chr4_0821*	[49]
pAOX176[b]	Methanol inducible	~26	/	/	[55]
pAOX737-△D+3D[c]	Methanol inducible	~28	/	/	[56]
pAOX2	Methanol inducible	1.8 ± 0.2	/	*PAS_chr4_0152*	[52]
pAOX2-mutant[d]	Methanol inducible	~18	/	/	[58]
pCAT1	Methanol inducible	~18	/	*PAS_chr2-2_0131*	[62]
pCAT1-mutant[e]	Methanol inducible	~23	/	*PAS_chr2-2_0131*	[62]
pDAS2	Methanol inducible	24.1 ± 3.3	/	*PAS_chr3_0834*	[52]
pFLD1	Methanol inducible	10.9 ± 1.1	/	*PAS_chr3_1028*	[52]
pPMP20	Methanol inducible	16.4 ± 3.3	/	*PAS_chr1-4_0547*	[52]
pDAS1	Methanol inducible	14.7 ± 2.7	/	*PAS_chr3_0832*	[52]
pFDH1	Methanol inducible	14.2 ± 2.2	/	*PAS_chr3_0932*	[52]
p0374	Methanol inducible	~4	/	*PAS_chr3_0374*	[49]
p0319	Methanol inducible	~3	/	*PAS_chr1-1_0319*	[49]
p0547	Methanol inducible	~2	/	*PAS_chr1-4_0547*	[49]
p0472[f]	Constitutive	~3 or 22	~0.5	*PAS_chr2-1_0472*	[49,54]
pGAP	Constitutive	1.0	~2.3	*PAS_chr2-1_0437*	[49]
p0769	Constitutive	~0.4	~1.8	*PAS_chr2-1_0769*	[49]
p0072	Constitutive	~0.3	~1.4	*PAS_chr1-1_0072*	[49]
pPDC[g]	Constitutive	~0.5	~3.5 or 4.7	*PAS_chr3_0188*	[49,53]
pPDI1	Constitutive	~0.5	~0.4	*PAS_chr4_0844*	[63]

Note.

[a] Promoter strengths are normalized to that of the constitutive promoter (*pGAP*) under methanol conditions.

[b] *pAOX176* is generated by removing a small number of bases before the TATA box of *pAOX1*.

[c] *pAOX737* represents that the bases from −940 to −737 bp of *pAOX1* are deleted. Region D is defined as the position of −638 to −510 bp in *pAOX1*. -△D+3D means the deletion of region D followed by the addition of 3 copies of region D at the 5′-end of *pAOX1*.

[d] *pAOX2*-mutant is obtained by mutating the bases at −255 and −456 positions of *pAOX2*.

[e] *pCAT1*-mutant represents that the bases from −433 to −411 bp of *pCAT1* are duplicated.

[f] *p0472* is 3 times stronger than *pGAP* when expressing recombinant α-amylase, while 22 times stronger when expressing xylanase.

[g] *pPDC* is 1.5 times stronger than *pGAP* when expressing recombinant α-amylase, while 2 times stronger when expressing human growth hormone.

promoters, existing promoters can be modified to possess the desirable features. For example, at least 12 cis-acting elements were found in the *AOX1* promoter and a promoter library of *pAOX1* is constructed by deleting and replicating the putative transcription factor binding modules. The activity range of the promoter library is 10%–160% of the wild-type *pAOX1* [55]. A key transcription factor, called Mxr1p, was

determined to be essential for the induction of *pAOX1* by methanol [56]. Due to the toxicity of methanol, Shen et al. established a methanol-free expression system using *AOX1* promoter mutants, in which the mutant promoter was induced by dihydroxyacetone and suppressed by glucose. The strength of this protein expression system can reach 50%–60% of the wild-type *pAOX1* expression system [57]. Dai et al. identified a *pAOX2* mutant, whose expression level was even higher than that of *pAOX1*. DNA sequencing revealed two point mutations at positions of −529 bp and −255 bp responsible for the dramatically improved promoter strength [58].

Terminators have been found to have important regulatory effects on transcription termination and the half-life of mRNA in *S. cerevisiae* [59, 60]. However, the significance of terminators is largely overlooked and little work has been done on the characterization of *P. pastoris* terminators. Vogl et al. tested the effect of different terminators on the expression of eGFP (enhanced green fluorescent reporter protein) under the control of *AOX1* promoter and *AOX1* terminator was found to enable the highest fluorescence intensity. In addition, inserting *Not*I restriction site into the *AOX1* terminator can further increase the fluorescence intensity by 37% [52]. Ito et al. characterized 72 terminators derived from *P. pastoris*, *S. cerevisiae*, and synthetic terminators, and found that the tunable range could reach up to 17-fold. Interestingly, the *S. cerevisiae* terminators could maintain their function after being transferred to *P. pastoris* [61]. These preliminary studies indicated the significance of terminators in regulating the expression level of heterologous genes and more mechanistic studies should be carried out in the near future.

2.3. Genome editing tools

As a fundamental tool, genome editing technology is essential for establishing *P. pastoris* as cell factories for recombinant proteins and value-added compounds. In the very beginning, site-directed gene integration and gene knockout were achieved through homologous recombination. Construction of a selection marker-containing plasmid that is capable of gene replacement in *P. pastoris* is one of the first genome editing tools [64]. For example, *HIS4*, *URA3,* and *URA5* genes are often used as selection markers in the corresponding defective *P. pastoris* strains [65]. However, these genome editing techniques usually leave selection marker expression cassettes in the host, which is not desirable for subsequent genetic manipulations and industrial applications. To enable multiple rounds of genome editing, Cre/*lox*P system was introduced into *P. pastoris*. Cre is a site-specific recombinase that specifically recognizes and recombines genes between two *lox*P loci. The advantage of this technology is that antibiotic resistance genes can be used for screening first and then recycled after the disruption of the target gene [66]. In addition, *mazF*, a toxic gene from *E. coli*, was used to construct a set of counter-selection techniques for marker-less genome editing in *P. pastoris* [67].

In recent years, emerging genome editing tools, such as ZFN (Zinc-finger nucleases), TALEN (transcription activator-like effector nucleases), and CRISPR/Cas, have revolutionized our capability of genetic manipulations of microbial cell factories (Fig. 2). These technologies use specific nucleases to create double-strand breaks (DSB) at the corresponding loci, which are repaired by homologous recombination (HR) or non-homologous end joining (NHEJ) to achieve the desirable genome editing. Particularly, the CRISPR/Cas system is the most widely used and most powerful genome editing technology. The CRISPR/Cas9 system is derived from the immune defense systems of bacteria and archaea [68], and has received in-depth research in microbial cell factories development, plant breeding, animal breeding, disease modeling, and biotherapy [69]. Weninger et al. systematically optimized the CRISPR/Cas9 expression system to achieve efficient and precise genome editing in *P. pastoris*, including but not limited to Cas9 coding sequences, gRNA sequences, gRNA structures (i.e. with ribozyme sequences), and
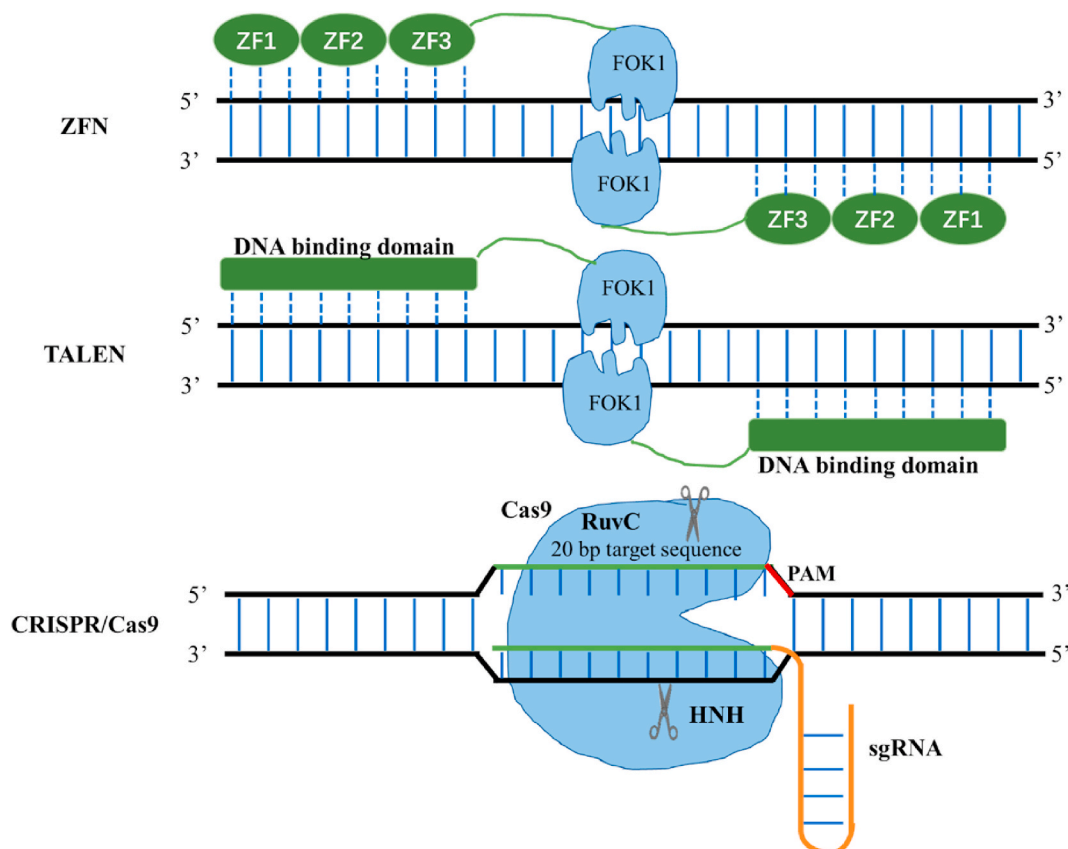

Fig. 2. Three emerging genome editing techniques, including ZFN, TALEN and CRISPR/Cas9.

promoters for the expression of Cas9 and gRNAs [70]. Among 95 combinations, only 6 constructs were found to be functional for genome editing, indicating the necessity for further optimization (Table 3). For example, Gu et al. found that the replacement of the origin of replication of the bearing plasmid from PARS1 to panARS increased the disruption efficiency of *ADE2* locus from ~10% to ~80% [32]. In addition, Dalvie et al. developed a sequencing-based strategy for the design of host-specific cassettes for modular and efficient expression of gRNAs and achieved high genome editing efficiency up to 95% [71]. Besides gene disruption, multiplex integration of heterologous genes is another essential synthetic biology tool for establishing *P. pastoris* as cell factories for natural products. Simultaneous integration of multiple genes was reported in a *KU70*-deficient *P. pastoris* strain, with an integration efficiency ranged from 57.7% to 70% and 12.5%–32.1% for double- and triple-loci, respectively [72].

3. Engineering of *P. pastoris* to produce natural products

3.1. Terpenoids

Terpenoids are value-added natural products derived from mevalonate and widely existed in nature, including but not limited to higher plants, fungi, and microorganisms. Many terpenoids have been found applications in medicine, food, cosmetics, animal feeds, and industry, leading to the exploration of the production of terpenoids using microbial cell factories. Bhataya et al. introduced the lycopene biosynthetic pathway into non-carotenogenic *P. pastoris* for the first time. Two lycopene-pathway plasmids were constructed, with plasmid pGAPZB-EBI* harboring genes *crtE*, *crtB*, and *crtI* and plasmid pGAPZB-EpBpI*p harboring the same set of genes with a peroxisomal targeting sequence (PTS1). Similar amount of lycopene was produced in the two yeast strains, indicating that the supply of FPP might be limited in *P. pastoris*. One clone expressing pGAPZB-EpBpI*p with the highest lycopene production was identified and further optimized by investigating the effects of culturing conditions (i.e. carbon sources and aerations). Finally, the production of lycopene reached up to 73.9 mg/L in the basic medium with glucose as the carbon source [77]. Later, β-carotene was synthesized by additionally integrating the lycopene β-cyclase gene from *Ficus carica* into the chromosome of the lycopene-producing strain, leading to the production of 339 μg of *β*-carotene per gram dry cell weight (DCW) [17]. Starting from the β-carotene-producing strain, further introduction of β-carotene ketolase gene (*crtW*) and β-carotene hydroxylase gene (*crtZ*) from *Agrobacterium aurantiacum* resulted in the production of 3.7 μg/g DCW of astaxanthin in *P. pastoris* [78]. In another study, Vogl et al. characterized a panel of promoters in the methanol utilization pathway of *P. pastoris*, which were further employed for combinatorial optimization of the β-carotene biosynthetic pathway. With different combinations of the methanol inducible promoters, the production of β-carotene can be varied for more than 10-fold. Via choosing appropriate promoters from the established promoter library, the yield of β-carotene reached up to 5 mg/g DCW [52].

(+)-Nootkatone, an excellent fragrance and insect repellent, have also been successfully produced in *P. pastoris*. The introduction of valencene synthase resulted in the biosynthesis of (+)-valencene. Followed by the co-expression of the premnaspirodiene oxygenase from *Hyoscyamus muticus* (HPO) and the cytochrome P450 reductase from *Arabidopsis thaliana*, (+)-valencene was hydroxylated to produce trans-nootkatol. Trans-nootkatol was then oxidized to (+)-nootkatone by the intrinsic activity of *P. pastoris*. The production of (+)-nootkatone was 17 mg/L in a shake flask and 208 mg/L in a bioreactor, respectively [19]. Interestingly, the overexpression of *RAD52*, which is responsible for DNA repair and recombination, improved the production of trans-nootkatol by 5-fold [79].

Dammarenediol-II is a triterpenoid with multiple pharmacological activities. On the basis of the natural triterpene biosynthesis pathway [80,81], Liu et al. introduced *PgDDS* from *Panax ginseng*, encoding a dammarenediol-II synthase that catalyzed the production of dammarenediol-II from 2,3-oxidosqualene, to successfully construct a dammarenediol-II producing *P. pastoris* strain (Fig. 3). By increasing the expression of *ERG1* to enhance the supply of 2,3-oxidosqualene and downregulating the expression of *ERG7* to decrease the production of lanosterol from 2,3-oxidosqualene, the yield of dammarenediol-II was increased from 0.03 mg/g DCW to 0.736 mg/g DCW. Finally, by extra supplementation of 0.5 g/L squalene into the culture medium, the yield of dammarenediol-II reached up to 1.073 mg/g DCW.

Similarly, Sun et al. established a menaquinone-4 (MK-4) *P. pastoris* cell factory by introducing a heterologous gene encoding *Homo sapiens* UBIAD1 (*HsUBIAD1*), which can produce MK-4 from phylloquinone (VK1) or menadione (VK3). *HsUBIAD1* was cloned into pGAPZA (with the constitutive promoter *pGAP*) and pPICZA (with the inducible promoter *pAOX1*) and the effect of promoters on the expression of the target gene was investigated. It was found that the vector pGAPZA (with the target gene *HsUBIAD1* under the control of *pGAP*) resulted in higher protein expression level. Then the geranylgeranyl pyrophosphate synthase gene (*GGPPS*) from *Sulfolobus acidocaldarius* was fused with the endogenous isopentenyl diphosphate isomerase gene (*IDI1*), and the resultant *IDI1-GGPPS* chimeric gene was integrated into the 28S ribosomal DNA (rDNA) loci in a multi-copy manner using a modified integrative vector (pGrG, based on pGAPZA). In combination with the optimization of the fermentation conditions (i.e. pH and temperature) resulted in the maximum yield of MK-4 up to 0.24 mg/g DCW [82].

Table 3
CRISPR/Cas9 systems for genome editing of *P. pastoris*.

Cas9 promoter	sgRNA promoter, promoter type	Host	Target(s)	Donor length	Efficiency	References
pHTX1	*pHTX1*, II	CBS7435	*GUT1*	1000 bp	87–94%	[70]
pENO1	*ptRNA-tRNA1*, III	NRRL Y-11430	*GUT1*	500 bp	95%	[71,73]
pHTX1	*pHTX1*, II	GS115 △*ku70*	2 loci[a]	1000 bp	57.7–70%	[72]
pHTX1	*pHTX1*, II	GS115 △*ku70*	3 loci[b]	1000 bp	12.5–32%	[72]
pGAP	*pHTX1*, II	GS115	*MXR1*	~600 bp	>80%	[74]
pGAP	*pSER*, III	GS115	*ADE2*	250 bp	80%	[32]
pGAP	*pHTX1*, II	GS115	*Gt1*	None	100%	[31]
pHTX1	*pHTX1*, II	CBS7435 △*ku70*	*GUT1*	1000 bp	78–91%[c]	[75]
pHTX1	*pHTX1*, II	CBS7435 △*ku70*	*GUT1*	1000 bp	80–95%[d]	[75]
pHTX1	*pHTX1*, II	CBS7435 △*ku70*	*GUT1*	1000 bp	100%[e]	[75]
pHTX1	*pHTX1*, II	KM71	*PDC1*	1000 bp	N.A	[76]

[a] Any two loci of *pAOX1*, *pFLD1,* and *pTEF1* were simultaneously targeted.

[b] *pAOX1*, *pFLD1,* and *pTEF1* were simultaneously targeted. None means that no donor was added and DSB was repaired by NHEJ during CRISPR editing.

[c] Integration efficiency using the canonically designed DNA donor.

[d] Integration efficiency using DNA donor with a selection marker (zeocin[r]).

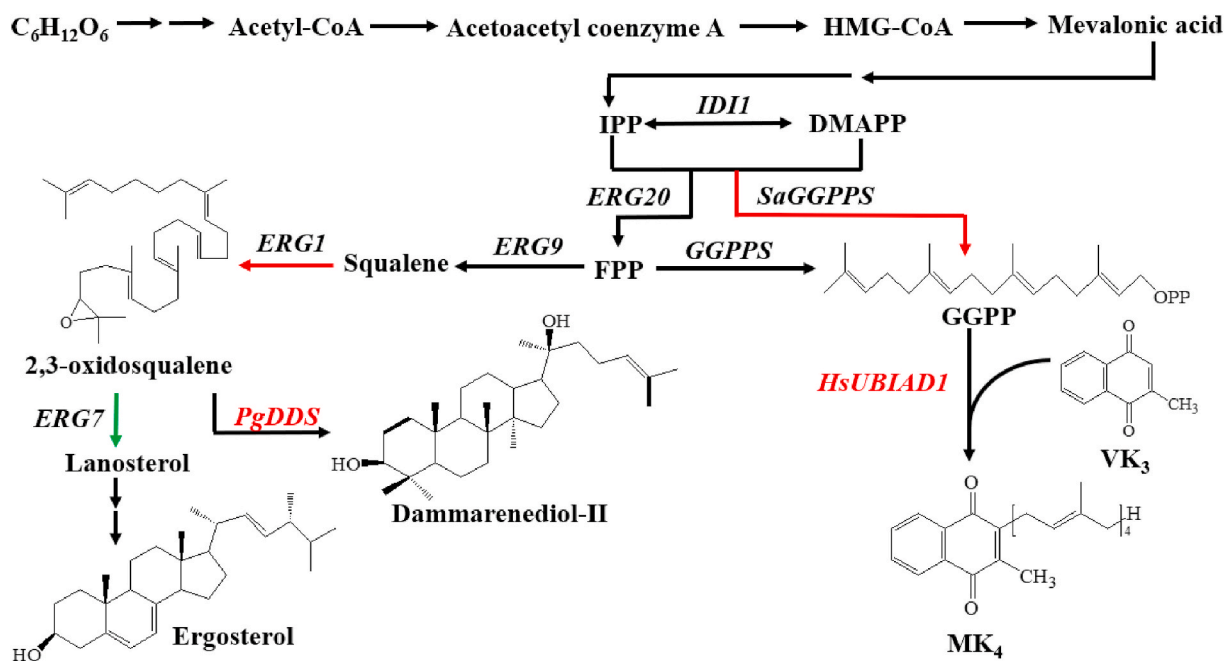[e] Integration efficiency using DNA donor with a replicon (ARS).

Fig. 3. Biosynthetic pathway of dammarenediol-II and MK_4 in *P. pastoris*. This study was carried out on the basis of the natural triterpene synthesis pathway. By introducing the exogenous *PgDDS* gene, encoding a dammarenediol synthase, the target compound dammarenediol-II was produced from 2,3-oxidosqualene in *P. pastoris*. The green arrow indicated the down-regulated gene (*ERG7*) and the red arrow indicated the overexpressed genes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.2. Polyketides

Polyketides are a class of secondary metabolites produced by bacteria, fungi, plants, and animals and the most important source of natural product-based drugs. 6-Methylsalicylic acid (6-MSA) is the first polyketide produced by *P. pastoris*. The 6-MSA biosynthetic pathway consisting of the phosphopantetheinyl transferase (PPtase) gene from *A. nidulans* and the 6-MSA synthase (6-MSAS) gene from *A. terrus* was successfully reconstituted in *P. pastoris*. After methanol induction, the production of 6-MSA was up to 2.2 g/L in 20 h in a 5 L bioreactor, which established *P. pastoris* as a promising cell factory for future industrial production of fungal polyketides [83].
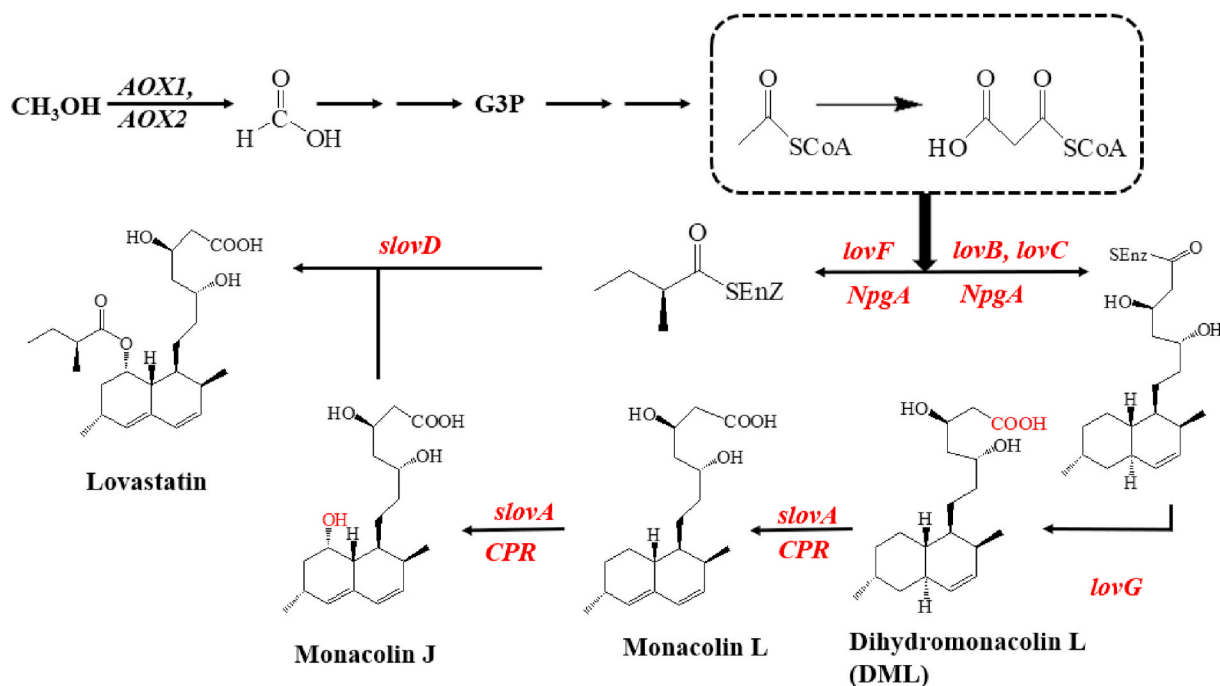

Fig. 4. Biosynthetic pathway for lovastatin and simvastatin. Heterologous genes integrated into the genome of *P. pastoris* were shown in red. *lovB* and *lovF*: two PKS genes; *lovC*: an enoyl-reductase gene; *lovG*: a thioesterase gene; *lovA*: a cytochrome P450 monooxygenase gene; and *lovD*: an acyl-transferase gene. *NpgA* is from *A. nidulans*. *lovB*, *lovC*, *lovF*, *lovG*, and *CPR* were amplified from the *A. terreus* genome. *slovA* and *slovD* are synthetic and codon-optimized DNA sequences for *lovA* and *lovD*, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Recently, *P. pastoris* was engineered for *de novo* biosynthesis of citrinin, a value-added compound. The structure and biosynthetic pathway of citrinin are more complicated than 6-MSA, serving as an excellent model compound for further investigations [84,85]. Besides the citrinin polyketide synthase gene *PksCT* (*CitS*) from *Monascus purpureus* and the phosphopantetheinyl transferase gene *NpgA* from *A. nidulans*, the citrinin gene cluster from *M. purpureus*, including a serine hydrolase gene *MPL1* (*CitA*), an oxygenase gene *MPL2* (*CitB*), a dehydrogenase gene *MPL4* (*CitD*), and other two intron-removed genes *MPL6* (*CitE*) and *MPL7* (*CitC*), was introduced to enable citrinin biosynthesis in *P. pastoris*. After 24 h induction with methanol, the yield of citrinin reached up to 0.6 ± 0.1 mg/L [86].

Production of monacolin J and lovastatin is another classic example of the production of polyketides using *P. pastoris* (Fig. 4). Seven enzymes, including lovastatin nonaketide synthase (LovB), enoyl reductase (LovC), thioesterase (LovG), a cytochrome P450 enzyme (LovA) together with a cytochrome P450 reductase (CPR), and cyltransferase (LovD) from *A. terreus*, as well as phosphopantetheinyl transferase (PPtase or NpgA) from *A. nidulans*, were heterologously expressed in *P. pastoris*. The expression of all these genes were driven by *pAOX1*. *LovB*, *LovC*, *LovG*, and *NpgA* expression cassettes were cloned into the vector pPICZ B. *LovA* and *CPR* were cloned into the vector pPIC3.5K. After these two vectors were linearized and integrated into the GS115 genome, the recombinant strains were screened using zeocin-containing plates and histidine-deficient plates, respectively. Under pH-controlled culture conditions with methanol as the carbon source, 60.0 mg/L of monacolin J and 14.4 mg/L of lovastatin were obtained. In order to overcome the limitations of intermediate accumulation and metabolic burden, approaches using pathway splitting at the branch point of dihydromonacolin L and co-cultivation of yeast species were developed. Under the optimal conditions, 593.9 mg/L monacolin J and 250.8 mg/L lovastatin were obtained [18]. Compared with the yield of lovastatin in *S. cerevisiae* (20 mg/L), *P. pastoris* obviously demonstrated higher potential for further development and practical applications [87].

3.3. Flavonoids

Flavonoids are widely existed in plants and used as essential components of many drugs, such as those to prevent cardiovascular and cerebrovascular diseases and treat chronic hepatitis [88]. Chang et al. used a recombinant *P. pastoris*, which expressed a fusion protein of CYP57B3 from *A. oryzae* and CPR from *S. cerevisiae*, to catalyze the *ortho*-hydroxylation of 10 flavonoids. The results showed that 5 flavonoids, including genistein, daidzein, liquiritigenin, naringenin, and apigenin could be transformed into the corresponding hydroxyl derivatives (Fig. 5). In addition, after feeding with genistein, the yeast extracts showed high inhibitory activity on melanogenesis [89]. Then, 3′-hydroxygenistein was identified as the active product of genistein biotransformation in recombinant *P. pastoris*, which demonstrated liver protection and anti-inflammatory effects. The conversion yield of 3′-hydroxygenisterin was 14% with a production level of 3.5 mg/L in a 5 L fermenter [90]. Furthermore, the periodic hydrogen peroxide shock strategy was employed to further increase the production of 3′-hydroxygenistein to 20.3 mg/L in a 5 L fermenter [91].

4. Conclusions and perspectives

Many synthetic biology tools have been developed to precisely control the expression of heterologous genes and the assembly and integration of multi-gene pathways in *P. pastoris* [75]. However, the development of *P. pastoris* cell factories for natural products is still limited to a few examples [82], particularly when compared with that of *S. cerevisiae*, indicating a need to develop novel synthetic biology tools. For example, the most commonly used promoters in *P. pastoris* are still *pAOX1* and *pGAP* [51], while the construction of efficient cell factories generally requires to precisely control the expression levels of biosynthetic pathway genes. In this case, promoters with different strength (weak, medium, and strong promoters) should be characterized under various fermentation conditions (i.e. carbon sources). In addition, most of exogenous genes are currently integrated into the *HIS4* or *AOX1* locus of the *P. pastoris* genome, while natural product biosynthetic pathways generally contain multiple genes [72]. In other words, well characterized integration sites are highly demanded to assemble natural product biosynthetic pathways. Ideally, the integration sites should enable efficient and stable expression of heterologous genes, without affecting cell fitness. Although CRISPR/Cas9 system has been established for genome
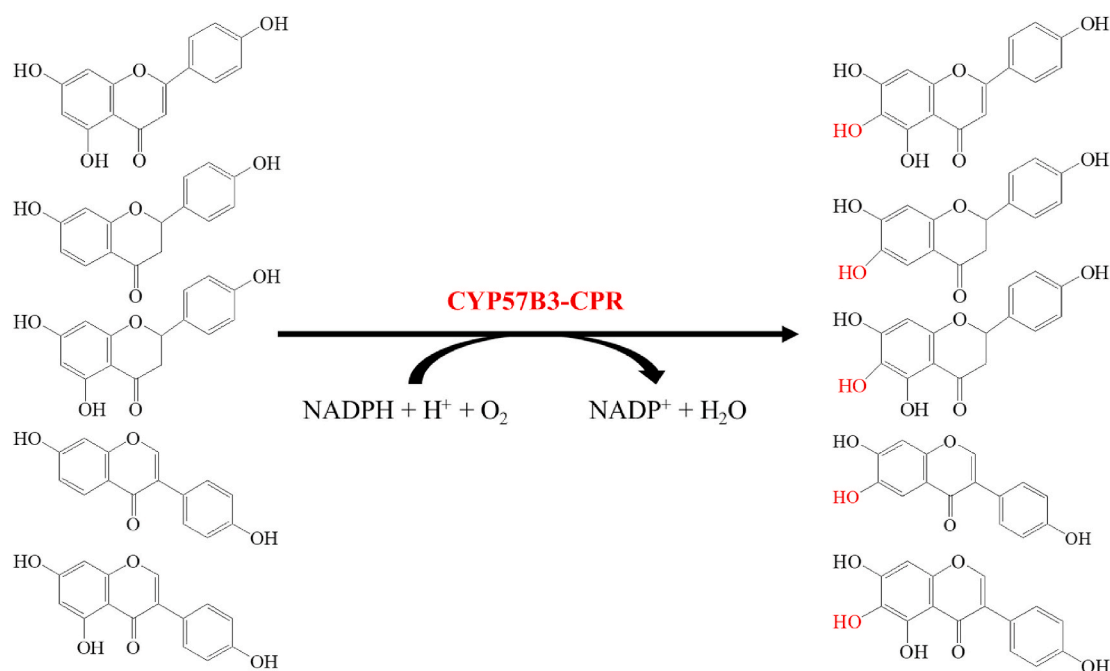

Fig. 5. The ortho-hydroxylation of five flavonoids. CYP57B3-CPR is a fusion protein consisting of CYP57B3 from *A. oryzae* and the cytochrome P450 reductase (CPR) from *S. cerevisiae*.

editing in *P. pastoris*, multiplex integration of long biosynthetic pathways still suffers from low efficiency and the reports on the use of CRISPR technology to construct *P. pastoris* cell factories are still rather limited. Currently, simultaneous integration of multiple genes was reported in a *ku70*-deficient *P. pastoris* strain. Unfortunately, the disruption of *KU70* was generally reported to result in impaired fitness and low transformation efficiency [72,92]. The challenge in improving homologous recombination efficiency without *KU70* disruption should be addressed for multiplex integration of natural product biosynthetic pathways in near future.

Due to the complexity of cellular metabolic network, systems level understanding is generally a prerequisite to establish efficient *P. pastoris* cell factories. Based on the genome sequencing results of *P. pastoris* DSMZ 70382 [93] and GS115 [94], genome-scale metabolic models (GEMs), PpaMBEL1254 [95], iPP668 [96], and iLC915 [97] have been established. Tomas-Gamisans et al. reconstructed and verified a new consensus model iMT1026 based on these three early reports of GEMs. New discoveries related to glycosylation, fatty acid metabolism, and cell energy were complemented to the new model. Growth rate, carbon dioxide production, arabitol production, and other parameters predicted by the model iMT1026 were consistent with the experimental results, confirming that the prediction and simulation capabilities have been improved [98]. The *P. pastoris* GEMs have been successfully implemented to identify targets to increase the production of recombinant proteins. Nocon et al. predicted 9 engineering targets based on GEMs, 5 of which significantly increased the production of cytosolic human superoxide dismutase (hSOD) [99]. Cankorur-Cetinkaya et al. used GEMs to analyze the metabolic burdens caused by heterologous protein synthesis and found that supplementation of tyrosine to the culture medium could increase the yield of human lysozyme and antibody fragment Fab-3H6 [100]. Currently, the production of natural products is still mainly limited by the low efficiency of the biosynthetic pathways, and there is a lack of examples on the application of GEMs in improving the production of natural products in *P. pastoris*. Nevertheless, once the bottleneck of the pathway enzymes has been addressed and the titer reaches to a certain level, the supply of the precursors and cofactors will become rate-limiting and GEMs can play a more important role in guiding the design of efficient yeast cell factories.

As a non-model yeast strain, our understanding of the metabolic and regulatory networks is still rather limited for *P. pastoris*. For example, Wriessnegger et al. found that the overexpression of *RAD52*, encoding a protein responsible for DNA repair and recombination, significantly improved the production of trans-nootkatol [79]. Therefore, it is highly desirable to develop genome-scale metabolic engineering strategies that can perturb all the genes at once, with an aim to identify non-intuitive engineering targets to improve the production of the desirable compounds and map genotype-phenotype relationships in a high-throughput manner [101]. Currently, CRISPR based genome-scale engineering has been well established in *E. coli* [102], *S. cerevisiae* [103–105], and mammalian cells [106]. Nevertheless, several challenges should be addressed to establish genome-scale engineering for *P. pastoris*, such as an efficient CRISPR system, high transformation efficiency, as well as high throughput screening.

In summary, great progress has been made in genetic manipulation (i.e. heterologous gene expression system and genome editing system) of *P. pastoris* for the production of recombinant proteins and value-added compounds. Nevertheless, more efforts should be devoted to developing novel synthetic biology tools (i.e. multi-gene pathway assembly and multiplex genome engineering) to establish *P. pastoris* as a robust chassis for the biosynthesis of natural products.

CRediT author statement

Juncan Gao and Lihong Jiang drafted the manuscript. Jiazhang Lian conceived the review idea and revised the manuscript. All authors approved the manuscript.

Declaration of competing interest

None.

Acknowledgements

References

[1] Rodrigues T, et al. Counting on natural products for drug design. Nat Chem 2016; 8:531–41.
[2] De Luca V, et al. Mining the biodiversity of plants: a revolution in the making. Science 2012;336:1658–61.
[3] Pena DA, et al. Metabolic engineering of *Pichia pastoris*. Metab Eng 2018;50:2–15.
[4] Schwarzhans JP, et al. Towards systems metabolic engineering in *Pichia pastoris*. Biotechnol Adv 2017;35:681–710.
[5] Lian J, Mishra S, Zhao H. Recent advances in metabolic engineering of *Saccharomyces cerevisiae*: new tools and their applications. Metab Eng 2018;50: 85–108.
[6] Jiang L, et al. Functional expression of eukaryotic cytochrome P450s in yeast. Biotechnol Bioeng 2020.
[7] Cregg JMCJ, Shi J, Higgins DR. Recombinant protein expression in *Pichia pastoris*. Mol Biotechnol 2000;23–52.
[8] Duan X, Gao J, Zhou YJ. Advances in engineering methylotrophic yeast for biosynthesis of valuable chemicals from methanol. Chin Chem Lett 2018;29: 681–6.
[9] Thak EJ, et al. Yeast synthetic biology for designed cell factories producing secretory recombinant proteins. FEMS Yeast Res 2020;20.
[10] Yang Z, Zhang Z. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: a review. Biotechnol Adv 2018;36:182–95.
[11] Carnicer M, et al. Development of quantitative metabolomics for *Pichia pastoris*. Metabolomics 2012;8:284–98.
[12] Zhou Q, et al. High-level production of a thermostable mutant of *Yarrowia lipolytica* lipase 2 in *Pichia pastoris*. Int J Mol Sci 2019;21.
[13] Yamada R, et al. Toward the construction of a technology platform for chemicals production from methanol: D-lactic acid production from methanol by an engineered yeast *Pichia pastoris*. World J Microbiol Biotechnol 2019;35:37.
[14] Yang Z, Zhang Z. Production of (2R, 3R)-2,3-butanediol using engineered *Pichia pastoris*: strain construction, characterization and fermentation. Biotechnol Biofuels 2018;11:35.
[15] Kong S, et al. De novo biosynthesis of 2-phenylethanol in engineered *Pichia pastoris*. Enzym Microb Technol 2020;133:109459.
[16] Siripong W, et al. Metabolic engineering of *Pichia pastoris* for production of isobutanol and isobutyl acetate. Biotechnol Biofuels 2018;11:1.
[17] Araya-Garay JM, et al. Construction of new *Pichia pastoris* X-33 strains for production of lycopene and beta-carotene. Appl Microbiol Biotechnol 2012;93: 2483–92.
[18] Liu Y, et al. Engineered monoculture and co-culture of methylotrophic yeast for *de novo* production of monacolin J and lovastatin from methanol. Metab Eng 2018;45:189–99.
[19] Wriessnegger T, et al. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. Metab Eng 2014;24:18–29.
20 Juergen N, Grantham N. In: Yeast strains for protein production. United States patent US; 2010. US2010/0279,348.
[21] Yoshimasu MA, et al. Soluble expression and purification of porcine pepsinogen from *Pichia pastoris*. Protein Expr Purif 2002;25:229–36.
[22] Ha SH, et al. Molecular cloning and high-level expression of G2 protein of hantaan (HTN) virus 76-118 strain in the yeast *Pichia pastoris* KM71. Virus Gene 2001;22:167–73.
[23] Wang Y, et al. Expression of *Aspergillus niger* glucose oxidase in *Pichia pastoris* and its antimicrobial activity against *Agrobacterium* and *Escherichia coli*. PeerJ 2020;8. e9010.
[24] Gaffar S, et al. Combination of genetic manipulation improved *Saccharomycopsis fibuligera* α-amylase secretion by *Pichia pastoris*. Indonesian Journal of Chemistry 2019;19.
[25] Khan MSI, et al. Molecular cloning and expression of recombinant *Trichoderma harzianum* chitinase in *Pichia pastoris*. Adv Life Sci 2020;7:122–8.
[26] Papakonstantinou T, Harris S, Hearn MT. Expression of GFP using *Pichia pastoris* vectors with zeocin or G-418 sulphate as the primary selectable marker. Yeast 2009;26:311–21.
[27] Marx H, et al. Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. FEMS Yeast Res 2009;9: 1260–70.
[28] Sunga AJ, Tolstorukov I, Cregg JM. Posttransformational vector amplification in the yeast *Pichia pastoris*. FEMS Yeast Res 2008;8:870–6.
[29] Aw R, Polizzi KM. Liquid PTVA: a faster and cheaper alternative for generating multi-copy clones in *Pichia pastoris*. Microb Cell Factories 2016;15:29.

[30] Li D, et al. A novel vector for construction of markerless multicopy overexpression transformants in *Pichia pastoris*. Front Microbiol 2017;8:1698.

[31] Yang Y, et al. High efficiency CRISPR/Cas9 genome editing system with an eliminable episomal sgRNA plasmid in *Pichia pastoris*. Enzym Microb Technol 2020:138.

[32] Gu Y, et al. 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;35:79.

[33] Camattari A, et al. Characterization of a panARS-based episomal vector in the methylotrophic yeast *Pichia pastoris* for recombinant protein production and synthetic biology applications. Microb Cell Factories 2016;15:139.

[34] Zhang T, et al. The expression of recombinant human LOX-1 and identifying its mimic ligands by fluorescence polarization-based high throughput screening. J Biotechnol 2006;125:492–502.

[35] Zhou HY, et al. Heterologous expression and biochemical characterization of a thermostable endo-beta-1,4-glucanase from *Colletotrichum orchidophilum*. Bioproc Biosyst Eng 2020.

[36] Wang D, Zheng P, Chen P. Production of a recombinant alpha-L-rhamnosidase from *Aspergillus niger* CCTCC M 2018240 in *Pichia pastoris*. Appl Biochem Biotechnol 2019;189:1020–37.

[37] Chen Z, et al. Recombinant antimicrobial peptide hPAB-beta expressed in *Pichia pastoris*, a potential agent active against methicillin-resistant *Staphylococcus aureus*. Appl Microbiol Biotechnol 2011;89:281–91.

[38] Deng H, et al. Characterisation of a monooxygenase in *Shiraia bambusicola*. Microbiology (Read) 2018;164:1180–8.

[39] Li X, et al. Overexpression of *Candida rugosa* lipase Lip1 via combined strategies in *Pichia pastoris*. Enzym Microb Technol 2016;82:115–24.

[40] Troshagina DS, et al. Cloning of phytase genes from *pantoea* sp. 3.5.1 and *Bacillus ginsengihumi* M2.11 in *Pichia pastoris*. BioNanoScience 2018;8:1045–53.

[41] Mombeni M, et al. pMOX: a new powerful promoter for recombinant protein production in yeast *Pichia pastoris*. Enzym Microb Technol 2020;139:109552.

[42] Du M, Battles MB, Nett JH. A color-based stable multi-copy integrant selection system for *Pichia pastoris* using the attenuated *ADE1* and *ADE2* genes as auxotrophic markers. Bioeng Bugs 2012;3:32–7.

[43] Yang J, et al. Hygromycin-resistance vectors for gene expression in *Pichia pastoris*. Yeast 2014;31:115–25.

[44] Schwarzhans JP, et al. A mitochondrial autonomously replicating sequence from *Pichia pastoris* for uniform high level recombinant protein production. Front Microbiol 2017;8:780.

[45] Chen H, et al. Engineering of an episomal plasmid suitable for high-throughput expression in *Pichia pastoris*. Comb Chem High Throughput Screen 2017;20:726–33.

[46] Lee CC, et al. An episomal expression vector for screening mutant gene libraries in *Pichia pastoris*. Plasmid 2005;54:80–5.

[47] Sasagawa T, et al. High-throughput recombinant gene expression systems in *Pichia pastoris* using newly developed plasmid vectors. Plasmid 2011;65:65–9.

[48] Schmidt A, et al. Thermostability improvement of *Aspergillus awamori* glucoamylase via directed evolution of its gene located on episomal expression vector in *Pichia pastoris* cells. Protein Eng Des Sel 2019;32:251–9.

[49] Xu N, et al. Identification and characterization of novel promoters for recombinant protein production in yeast *Pichia pastoris*. Yeast 2018;35:379–85.

[50] Yurimoto H, Oku M, Sakai Y. Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis. Internet J Microbiol 2011;2011:101298.

[51] Zhang AL, et al. Recent advances on the GAP promoter derived expression system of *Pichia pastoris*. Mol Biol Rep 2009;36:1611–9.

[52] Vogl T, et al. A toolbox of diverse promoters related to methanol utilization: functionally verified parts for heterologous pathway expression in *Pichia pastoris*. ACS Synth Biol 2016;5:172–86.

[53] Massahi A, Çalık P. Naturally occurring novel promoters around pyruvate branch-point for recombinant protein production in *Pichia pastoris* (*Komagataella phaffii*): pyruvate decarboxylase- and pyruvate kinase- promoters. Biochem Eng J 2018;138:111–20.

[54] Karaoglan M, Karaoglan FE, Inan M. Comparison of *ADH3* promoter with commonly used promoters for recombinant protein production in *Pichia pastoris*. Protein Expr Purif 2016;121:112–7.

[55] Hartner FS, et al. Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. Nucleic Acids Res 2008;36:e76.

[56] Xuan Y, et al. An upstream activation sequence controls the expression of *AOX1* gene in *Pichia pastoris*. FEMS Yeast Res 2009;9:1271–82.

[57] Shen W, et al. A novel methanol-free *Pichia pastoris* system for recombinant protein expression. Microb Cell Factories 2016;15:178.

[58] Dai XY, et al. Isolation and characterization of PAOX2 mutant in *Pichia pastoris*. Yi Chuan Xue Bao 2000;27:641–6.

[59] Curran KA, et al. Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. Metab Eng 2013;19:88–97.

[60] Sheng Y, Zhang G. Yeast terminator engineering: from mechanism exploration to artificial design. Synthetic Biology Journal 2020;1:709–21.

[61] Ito Y, et al. Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. Nucleic Acids Res 2020;48:13000–12.

[62] Nong L, et al. Engineering the regulatory site of the catalase promoter for improved heterologous protein production in *Pichia pastoris*. Biotechnol Lett 2020.

[63] Prattipati M, Ramakrishnan K, Sankaranarayanan M. *Pichia pastoris* Protein Disulfide Isomerase (PDI1) promoter for heterologous protein production and its sequence characterization. Enzym Microb Technol 2020;140:109633.

[64] Soderholm J, Bevis BJ, Glick BS. Vector for pop-in/pop-out gene replacement in *Pichia pastoris*. Biotechniques 2001;31:306–10. 312.

[65] Nett JH, Gerngross TU. Cloning and disruption of the *PpURA5* gene and construction of a set of integration vectors for the stable genetic modification of *Pichia pastoris*. Yeast 2003;20:1279–90.

[66] Piva LC, et al. Acetamidase as a dominant recyclable marker for *Komagataella phaffii* strain engineering. Appl Microbiol Biotechnol 2018;102:2753–61.

[67] Yang J, Jiang W, Yang S. *mazF* as a counter-selectable marker for unmarked genetic modification of *Pichia pastoris*. FEMS Yeast Res 2009;9:600–9.

[68] Barrangou R. The roles of CRISPR–Cas systems in adaptive immunity and beyond. Curr Opin Immunol 2015;32:36–41.

[69] Chen B, et al. Recent advances in CRISPR research. Protein Cell 2020.

[70] Weninger A, et al. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. J Biotechnol 2016;235:139–49.

[71] Dalvie NC, et al. Host-informed expression of CRISPR guide RNA for genomic engineering in *Komagataella phaffii*. ACS Synth Biol 2019.

[72] Liu Q, et al. CRISPR-Cas9-mediated genomic multiloci integration in *Pichia pastoris*. Microb Cell Factories 2019;18:144.

[73] Brady JR, et al. Identifying improved sites for heterologous gene integration using ATAC-seq. ACS Synth Biol 2020;9:2515–24.

[74] Hou C, et al. Targeted editing of transcriptional activator *MXR1* on the *Pichia pastoris* genome using CRISPR/Cas9 technology. Yeast 2020;37:305–12.

[75] Weninger A, et al. Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers. J Cell Biochem 2018;119:3183–98.

[76] Siripong W, et al. Metabolic engineering of *Pichia pastoris* for production of isopentanol (3-Methyl-1-butanol). Enzym Microb Technol 2020;138:109557.

[77] Bhataya A, Schmidt-Dannert C, Lee PC. Metabolic engineering of *Pichia pastoris* X-33 for lycopene production. Process Biochem 2009;44:1095–102.

[78] Araya-Garay JM, et al. Construction of a novel *Pichia pastoris* strain for production of xanthophylls. Amb Express 2012;2:24.

[79] Wriessnegger T, et al. Enhancing cytochrome P450-mediated conversions in *P. pastoris* through *RAD52* over-expression and optimizing the cultivation conditions. Fungal Genet Biol 2016;89:114–25.

[80] Chu LL, Montecillo JAV, Bae H. Recent advances in the metabolic engineering of yeasts for ginsenoside biosynthesis. Front Bioeng Biotechnol 2020;8:139.

[81] Liu XB, et al. Metabolic engineering of *Pichia pastoris* for the production of dammarenediol-II. J Biotechnol 2015;216:47–55.

[82] Sun X, et al. Construction of a novel MK-4 biosynthetic pathway in *Pichia pastoris* through heterologous expression of *Hs*UBIAD1. Microb Cell Factories 2019;18:169.

[83] Gao L, et al. Engineered fungal polyketide biosynthesis in *Pichia pastoris*: a potential excellent host for polyketide production. Microb Cell Factories 2013;12:77.

[84] Sakai K, et al. Construction of a citrinin gene cluster expression system in heterologous *Aspergillus oryzae*. J Biosci Bioeng 2008;106:466–72.

[85] Shimizu T, Kinoshita H, Nihira T. Identification and in vivo functional analysis by gene disruption of *ctnA*, an activator gene involved in citrinin biosynthesis in *Monascus purpureus*. Appl Environ Microbiol 2007;73:5097–103.

[86] Xue Y, et al. Methylotrophic yeast *Pichia pastoris* as a chassis organism for polyketide synthesis via the full citrinin biosynthetic pathway. J Biotechnol 2017;242:64–72.

[87] Xu W, et al. LovG: the thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. Angew Chem Int Ed Engl 2013;52:6472–5.

[88] Li Y-J, et al. Screening and characterization of natural antioxidants in four *Glycyrrhiza* species by liquid chromatography coupled with electrospray ionization quadrupole time-of-flight tandem mass spectrometry. J Chromatogr A 2011;1218:8181–91.

[89] Chang TS, Tsai YH. Inhibition of melanogenesis by yeast extracts from cultivations of recombinant *Pichia pastoris* catalyzing ortho-hydroxylation of flavonoids. Curr Pharmaceut Biotechnol 2015;16:1085–93.

[90] Ding H-Y, et al. Identification of 3′-hydroxygenistein as a potent melanogenesis inhibitor from biotransformation of genistein by recombinant *Pichia pastoris*. Process Biochem 2015;50:1614–7.

[91] Wang TY, et al. Improving 3'-hydroxygenistein production in recombinant *Pichia pastoris* using periodic hydrogen peroxide-shocking strategy. J Microbiol Biotechnol 2016;26:498–502.

[92] Ito Y, et al. Deletion of DNA ligase IV homolog confers higher gene targeting efficiency on homologous recombination in *Komagataella phaffii*. FEMS Yeast Res 2018;18.

[93] Mattanovich D, et al. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. Microb Cell Factories 2009;8:29.

[94] De Schutter K, et al. Genome sequence of the recombinant protein production host *Pichia pastoris*. Nat Biotechnol 2009;27:561–6.

[95] Sohn SB, et al. Genome-scale metabolic model of methylotrophic yeast *Pichia pastoris* and its use for in silico analysis of heterologous protein production. Biotechnol J 2010;5:705–15.

[96] Chung BK, et al. Genome-scale metabolic reconstruction and in silico analysis of methylotrophic yeast *Pichia pastoris* for strain improvement. Microb Cell Factories 2010;9:50.

[97] Caspeta L, et al. Genome-scale metabolic reconstructions of Pichia stipitis and *Pichia pastoris* and in silico evaluation of their potentials. BMC Syst Biol 2012;6: 24.

[98] Tomas-Gamisans M, Ferrer P, Albiol J. Integration and validation of the genome-scale metabolic models of *Pichia pastoris*: a comprehensive update of protein glycosylation pathways, lipid and energy metabolism. PloS One 2016;11. e0148031.

[99] Nocon J, et al. Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. Metab Eng 2014;24:129–38.

[100] Cankorur-Cetinkaya A, et al. Process development for the continuous production of heterologous proteins by the industrial yeast. *Komagataella phaffii*. Biotechnol Bioeng 2018;115:2962–73.

[101] Xia S, et al. Advances in genome evolution of *Saccharomyces cerevisiae*. Synthetic Biology Journal 2020;1:556–69.

[102] Garst AD, et al. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. Nat Biotechnol 2017; 35:48–55.

[103] Bao Z, et al. Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision. Nat Biotechnol 2018;36:505–8.

[104] Lian J, et al. Multi-functional genome-wide CRISPR system for high throughput genotype-phenotype mapping. Nat Commun 2019;10:5794.

[105] Lian J, et al. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. Nat Commun 2017;8:1688.

[106] Gilbert LA, et al. Genome-scale CRISPR-mediated control of gene repression and activation. Cell 2014;159:647–61.