frontiers | Research Topics

Efficient biomanufacturing via microbial cell factories, volume II

Edited by
Wei Luo, Zhenlin Han, Shuang Li
and Yinjie Tang

Published in
Frontiers in Bioengineering and Biotechnology

ISSN 1664-8714
ISBN 978-2-8325-3453-3
DOI 10.3389/978-2-8325-3453-3

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Efficient biomanufacturing via microbial cell factories, volume II

Topic editors

Wei Luo — Jiangnan University, China

Zhenlin Han — University of Hawaii at Manoa, United States

Shuang Li — South China University of Technology, China

Yinjie Tang — Washington University in St. Louis, United States

Citation

Luo, W., Han, Z., Li, S., Tang, Y., eds. (2023). *Efficient biomanufacturing via microbial cell factories, volume II*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-3453-3

Table of contents

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fbioe.2022.1010931

OPEN ACCESS

EDITED BY
Yuan Lu,
Tsinghua University, China

REVIEWED BY
Jin Huang,
Zhejiang University of Technology, China
Li-Hai Fan,
Fuzhou University, China

*CORRESPONDENCE
Wei Luo,
wluo@jiangnan.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 03 August 2022
ACCEPTED 15 August 2022
PUBLISHED 09 September 2022

CITATION
Yang Q, Cai D, Chen W, Chen H and Luo W (2022), Combined metabolic analyses for the biosynthesis pathway of L-threonine in *Escherichia coli*. *Front. Bioeng. Biotechnol.* 10:1010931. doi: 10.3389/fbioe.2022.1010931

Combined metabolic analyses for the biosynthesis pathway of L-threonine in *Escherichia coli*

Qiang Yang[1], Dongbo Cai[2], Wenshou Chen[2], Huiying Chen[3] and Wei Luo[1]*

[1]The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China, [2]State Key Laboratory of Biocatalysis and Enzyme Engineering, Environmental Microbial Technology Center of Hubei Province, College of Life Sciences, Hubei University, Wuhan, China, [3]College of Chemistry and Bioengineering, Guilin University of Technology, Guilin, China

Currently, industrial production of L-threonine (Thr) is based on direct fermentation with microorganisms such as *Escherichia coli*, which has the characteristics of low cost and high productivity. In order to elucidate the key metabolic features of the synthesis pathway of Thr in *E. coli* to provide clues for metabolic regulation or engineering of the strain, this study was carried out on an L-threonine over-producing strain, in terms of analyses of metabolic flux, enzyme control and metabonomics. Since environmental disturbance and genetic modification are considered to be two important methods of metabolic analysis, addition of phosphate in the media and comparison of strains with different genotypes were selected as the two candidates due to their significant influence in the biosynthesis of Thr. Some important targets including key nodes, enzymes and biomarkers were identified, which may provide target sites for rational design through engineering the Thrproducing strain. Finally, metabolic regulation aimed at one biomarker identified in this study was set as an example, which confirms that combined metabolic analyses may guide to improve the production of threonine in *E. coli*.

KEYWORDS

Escherichia coli, l-threonine, combined metabolic analyses, metabolic flow, enzyme control analysis, metabolomics

Introduction

L-threonine (Thr) is one of the eight essential amino acids and mainly used in food fortification, pharmaceuticals, chemical reagents and feed additives (Komatsubara et al., 1978; Follettie et al., 1988; Leuchtenberger et al., 2005), with the amount used in feed additives in particular growing rapidly (Zheng et al., 2020). It is often added to the feed of immature piglets and poultry and is the second limiting amino acid in pig feed and the third limiting amino acid in poultry feed. The current methods of producing Thr mainly include hydrolysis of animal proteins and microbial fermentation. Compared with the former, the microbial fermentation method has the advantages of low cost and low pollution (Li et al., 2021;). Among the Thr-producing strains, *E. coli* is the main host,

which is ascribed to its simple genetic background and convenience to modify through metabolic engineering (Debabov et al., 1981; Song et al., 2000). In addition, it is of short growth cycle, high cell intensity and low requirement for equipments. Efficient synthesis of Thr does not only depend on a large increase in the efficiency of a particular rate-limiting reaction, but requires a balance of multiple metabolic pathways in the biosynthetic network, such as intensification of the target metabolite flow, weakening the competitive branch metabolite flow, improvement of the extracellular transport efficiency, etc. (Dong et al., 2011; Lee et al., 2009).

Metabolic flux analysis (MFA) is a metabolic network analysis method via the intracellular reaction stoichiometry model, which is based on the pseudo-steady state assumption that the metabolic flux distribution of all metabolites is estimated in a dynamic equilibrium between the rates of production and consumption in the metabolic pathway (Lee et al., 2003; Park and Lee, 2010). This method has been successfully applied to the metabolic analyses of some amino acids and has resulted in significant savings in experimental costs. Since metabolite synthesis in the metabolic pathway is controlled by various enzyme-catalyzed reactions, it is necessary to investigate the control effect of these enzymes on carbon flux. Metabolic control theory believes that the change of metabolic pathway reaction at each step will cause the change of system parameters, so there is no stable rate-limiting step for biochemical reaction in metabolic pathway (Toya et al., 2011; Wang et al., 2014; Zhu et al., 2021). Flux control coefficient (FCC) and elasticity coefficient constitute the genetic theory of metabolic control analysis, which describe the global parameters in the whole metabolic network, and are also the bridge between enzymatic activities and metabolic fluxes. Another method for cell metabolic analysis relies on metabonomic technology, which detects the change of metabolites through GC/MS or LC/MS and accurately reflects the metabolic change under different conditions (Yang et al., 2018; Park et al., 2021). Metabonomics research now penetrated into many fields, such as medicine (Schmid et al., 2004; Antoniewicz, 2021), food (Picone et al., 2011; Antoniewicz, 2021) and microbiology (Luo et al., 2020).

In this study, phosphate was observed as a significant factor affecting the biosynthesis of Thr (Liu et al., 2020), thus was selected as an environment disturber to elucidate the biosynthesis pathway by metabolic flux analysis. The effect of phosphate on the metabolic flux distribution of Thr biosynthesis was investigated under fed-batch mode. On the other hand, the enzyme activities involved in Thr synthesis were measured and the metabolic control coefficients on fluxes were calculated by multivariate statistical analysis, which may provide a theoretical basis for the modification and metabolic regulation of key enzymes. Furthermore, the influence of genetic modification on the synthesis pathway of Thr through metabonomic analysis was investigated. The differential metabolites and metabolic pathways were observed in two strains with different genotypes and some biomarkers were obtained through multivariate statistical methods. Then metabolic regulation on a target obtained from the above metabolic pathway analyses was set as an example to improve the production of Thr.

Materials and methods

Strains, reagents and instruments

Thr over-producing srain *E. coli* TWF001 was used for the metabolic analyses for the biosynthesis pathway of Thr and W3110 was used as a control (Zhao et al., 2018). Methanol and acetonitrile with chromatographic grade were used as mobile phase in the determination by high performance liquid chromatography (HPLC). Pre-column derivatizer o-phthalaldehyde, sodium hydroxide, ammonia and 10% trichloroacetic acid were purchased from local market. Biosensor SBA-40C (Shandong Academy of Sciences, China) and Agilent HPLC were used for determination of glucose and other metabolites, respectively.

Culture media

Basic seed medium contained 10 g L^{-1} peptone, 5 g L^{-1} yeast extract and 10 g L^{-1} sodium chloride, which was adjusted to pH 7.0 and sterilised at 115°C for 30 min. Initial fermentation media (Zhao et al., 2020) contained 20 g L^{-1} glucose, 6.4 g L^{-1} beet molasses, 3 g L^{-1} maize pulp, 0.7 g L^{-1} betaine HCl, 0.4 g L^{-1} $MgSO_4$, 0.9 g L^{-1}KCl, 0.011 g L^{-1} $MnSO_4{\cdot}H_2O$, 0.011 g L^{-1} $FeSO_4{\cdot}7H_2O$ and 0.9 g L^{-1} H_3PO_4 at pH 7.0.

Seed culture and L-threonine fermentation

The conserved strains from the tube were picked out, streaked onto the prepared activation medium slant, and incubated at 37°C for 8–10 h a 250 ml flask supplemented with 50 ml of seed medium, which was incubated at 37 °C for 14 h in a shaker with 200 r·min^{-1}. For batch fermentation, 250 ml shake flasks supplemented with 50 ml of fermentation medium were inoculated with seed culture at 1% pitching rate and incubated at 37°C for 48 h in a shaker with 200 r min^{-1}.

Fed-batch fermentation of Thr was conducted at 37°C in a 5 L fermentor by continuous feeding 500 g L^{-1} glucose to maintain its concentration between 5 and 20 g L^{-1}. Dissolved oxygen was adjusted by ventilation rate and agitation speed at around 30% of the solubility saturation and pH was maintained at between 6.8 and 7.2 with 25% ammonia (Wang et al., 2014).

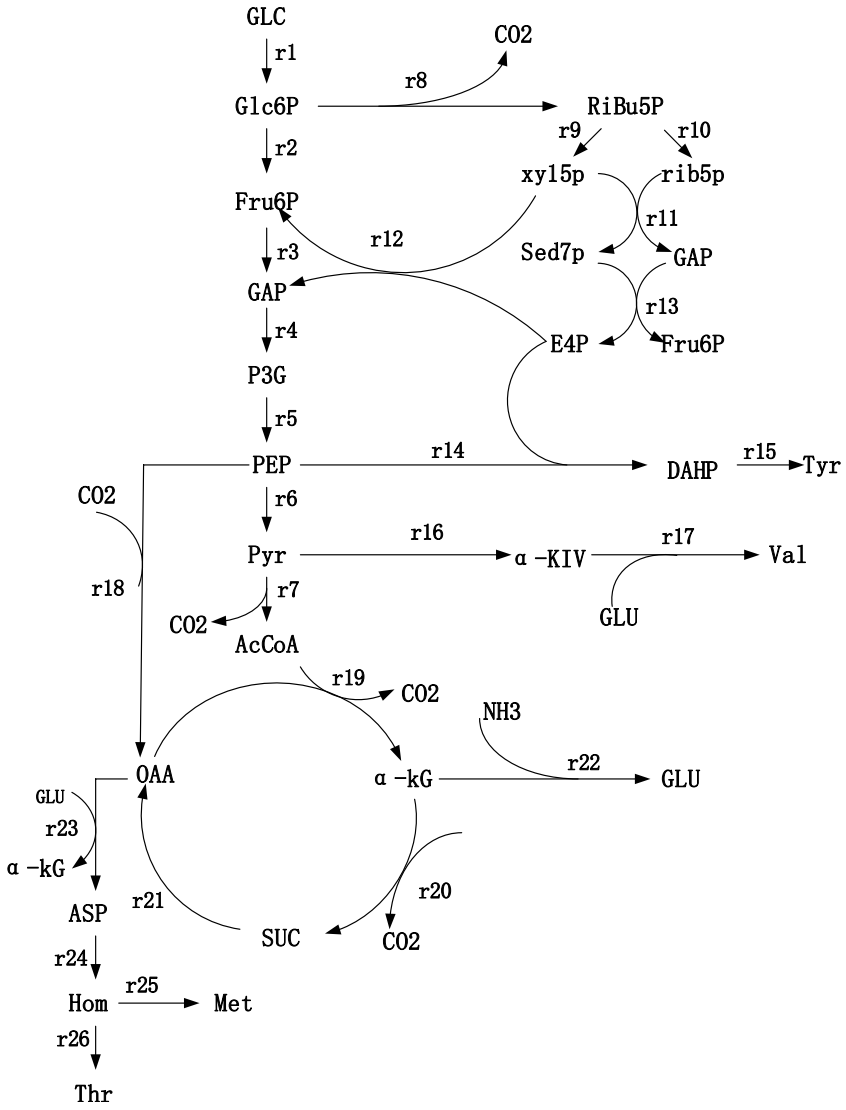


FIGURE 1
Metabolic network of Thr synthesis. 1) Glc: Glucose; 2) Glc6P: Glucose-6-phosphate; 3) PEP: Phosphoenolpyruvate; 4) Pyr: Pyruvate; 5) Fru6P: Fructose- 6-phosphate; 6) GAP: Glyceradehyde-3-phosphate; 7) P3G: 3-phosphoglycerate; 8) AcCoA: AcetylcoenzymeA; 9) Ribu5P: Ribulose-5-phosphate; (10) Xyl5P: Xylulose-5-phosphate; 11) Rib5P: Ribose-5-phosphate; 12) Sed7P: Sedoheptulose-7-phosphate; 13) E4P: Erythrose-4-phosphate; 14) OAA: Qxaloacetate; 15) α-KG: α-ketoglutarate; 16) NADPH: Nicotinamide adenine dinu-cleotide phosphate; 17) Glu: Glutamate; 18) αKiv: α-ketoisovalerate; 19) Asp: Aspartate; (20) Hom: Homoserine; 21) DAHP: 3-deoxy-d-arabi-noheptulosonate-7-phosphate.

Metabolic flow analysis

According to the literatures (Park and Lee, 2010; Su et al., 2018), Embden-Meyerhof-Parnas (EMP), tricarboxylic acid (TCA), hexose monophosphate pathway (HMP), anaplerotic route and phosphotransferase system (PTS) are present in *E. coli*. HMP pathway is important for amino acid synthesis since large amounts of reducing substances NADH/NADPH are produced to maintain cytoplasmic redox equilibrium. In addition, the glyoxalate cycle does not occur in *E. coli* when glucose is used as the substrate, indicating that the TCA cycle is still the main oxidative pathway in *E. coli*. Therefore, pathways of EMP, HMP and TCA were set as the main network pathways when constructing the Thr metabolic network.

According to Thr synthesis pathway in the KEGG database, some references (Schmid et al., 2004; OlafKrömera et al., 2006) and the metabolites detected in strains TWF001, the metabolic network was established based on the following principles, 1) Cells in the period of pseudo-steady state are of non-growth and the biomass changes can be ignored; 2) the total amount of NADPH consumed in the reaction pathway is equal to the total amount of NADPH produced by the TCA cycle and the HMP

pathway, i.e. the balance of NADPH supply and demand; 3) the glyoxalate cycle does not exist during cell metabolism; 4) reactions that proceed in a fixed ratio and intermediate reactions without branching points are simplified to a reaction equation; 5) during the stagnation phase of cell growth, the total cellular maintenance energy is not equal to the ATP consumption due to the presence of a large number of invalid cycles, so the balance of total ATP is not considered; the network of Thr biosynthesis metabolism is shown in Figure 1.

According to the assumption that the intermediate metabolites is in the pseudo-steady state, the change rate of intracellular metabolites is 0. The accumulation rate of metabolites is calculated according to the following equation,

$$r_i(t) = \sum_{i=1}^{M} \partial_{ij} r_j(t) - \sum_{i=1}^{i=N} \partial_{ik} r_k(t)$$

In this rate equation, $r_i(t)$ is the rate of accumulation of intermediate metabolite i (mmol·(Lh)$^{-1}$); $r_j(t)$ is the reaction rate of the reaction step j for the synthesis of i (mmol·(Lh)$^{-1}$]; $r_k(t)$ is the reaction rate of the reaction step k for the synthesis of i (mmol·(Lh)$^{-1}$); ∂_{ij} is the stoichiometry coefficient of the reaction step j; ∂_{ik} is the stoichiometry coefficient of the reaction step k. The rate equations at the metabolite nodes in the metabolic network were shown in Supplementary Table S2. The set of rate equations constructed consists of 21 equations with 26 unknowns and a degree of freedom of 5. The contents of glucose, Thr, L-tyrosine (Tyr), L-valine (Val) and L-methionine (Met) were measured, which were used as known parameters and substituted into the above set of metabolic rate equations to obtain the metabolic flow distribution using lingo software.

Determination of enzyme activities

When the cell growth entered stationary phase (under pseudo-steady state), samples were collected and centrifuged at 10,000 rpm for 5 min in a refrigerated centrifuge at 4°C. 0.5 g wet weight of the cells was washed by 10 ml of sterile 0.01 M PBS phosphate buffer for twice. Then 10 ml enzyme extraction solution were added for sonication and fragmentation treatment was under the following conditions, sonication power (PW) of 80%, working time of 2 s, intervals of 1 s, repeated 30 times. The samples were then centrifuged at 4°C for 5 min at 8,000 rpm to remove the precipitates and the enzyme solution was kept at a low temperature before determination of enzyme activities. The activities of pyruvate kinase (PK), malate dehydrogenase (MDH), fructose 1,6-diphosphate aldolase (FBA), glucose-6-phosphate dehydrogenase (G6PD), phosphoenolpyruvate carboxylase (PEPC) and hexokinase (HK) were determined by kits purchased from Solarbio Biotechnology Ltd. Protein concentrations were measured using the Coomassie Brilliant Blue G250 staining method at 595 nm and protein content was detected according to the increase in peak absorption.

Principal component analysis

For correlation analysis of the activities of different enzymes in the Thr synthesis pathway, the complex, multidimensional nature of the data requires the use of a multivariate data analysis method, i.e., principal component analysis (Antoniewicz, 2021), when the correlation between variables is significant. The six enzyme activities in the Thr synthesis pathway were measured under fed-batch mode and submitted to principal component analysis, which were set as initial indicators and noted as in order, X1 to X6. The screening conditions for principal component analysis were set as follows, the eigenvalues $\lambda > 1$, the principal components cumulatively reflecting more than 85% of the original information. Using SPSS25.0 software, the equation $F_j = a_{1j}X_1+a_{2j}X_2.....+a_{10j}X_{10}$ (j = 1, 2, 3.....10) was obtained, where the coefficients of the expression of the principal component reflect the combined influence of each enzyme on the principal (Vo et al., 2018). This analysis excludes the effect of enzyme synergy on Thr flux, as the principal component eliminates the correlation of the initial indicators.

Metabolism quenching and metabolites pre-treatment

Samples were collected and centrifugated at 4°C, 10,000 rpm for 5 min to remove the precipitates, and the supernatant was mixed with cold methanol with a volume ratio of 1:3. The mixes were shaked and precipitates occurred during this process were removed by centrifugation at 4°C, 10,000 rpm for 5 min. All samples were kept in an ice box throughout the operation and this process was repeated twice.

The extracted metabolites were quantified using an internal standard method (Luo et al., 2020) and the detailed operation steps are as follows. 500 μL of the extract were added with 30 μL of the internal standard reagent ribitol (1 mg mL^{-1}), and the mixes were shaken well and blown dry in a nitrogen blower. Then derivatization of metabolites was conducted by reactions of oximation and silylation. For the former, the samples were mixed with 350 μL of pyridinium methoxide hydrochloride solution (20 mg mL^{-1}) for 2 h at 30°C. Afterwards, 350 μL of BSTFA-TMCS (99:1, V/V) solution were added into the solution and the reaction was carried out at 70°C for 1 h. Then samples were blown dry with a nitrogen blower, and 1 ml of hexane/dichloroacetic acid was added to re-dissolve them. The samples were treated with a 0.22 μm filter membrane and detected on a gas chromatograph-mass spectrometer (GC-MS) machine (Jonsson et al., 2004; Koek et al., 2010).

Gas chromatograph-mass spectrometer condition

Metabolomic detection was conducted on Thermo GC-MS (TSQ8000; Thermo Scientific, United States) with a column (TG-

5, 30*0.25 mm*0.25 μm), using a helium carrier gas at a flow rate of 1.2 ml min^{-1} and operating at splitting ratio of 5.0. Electron energy and emission current were set to 70 eV and 25 μA, respectively. GC oven temperature was raised to 65°C, maintained for 2 min and then increased to 280°C at a constant speed in 13 min. Interface, transmission line and ion-source were kept at 300°C, 280°C and 300°C, respectively. The mass spectral scan range was set at 35–650 mz.

Data analyses

The metabolites were qualitatively analyzed by searching the spectrum in the GC-MS database or comparison with standards, and the relative contents of metabolites were determined according to the peak area of the internal standard ribitol in the solution. The data obtained were collated into csv format and imported into SIMCA-P software for principal component analysis and discriminant analysis. Cluster analysis made by Origin software was used to resolve intra- and inter-group relationship.

Metabolite determination by high performance liquid chromatography and biosensor

Samples were centrifuged at 10,000 rpm for 2 min to remove the precipitates and then treated with 10% trichloroacetic acid at equal volume for 4–8 h to remove proteins from the samples.

The determination of Thr and other amino acids was carried out by HPLC with a C18 column (5 μm, 250 mm × 4.6 mm) (Luo et al., 2021). The mobile phase was divided into two parts, A and B. The mobile phase A contained 10 mmoL·L^{-1} disodium hydrogen phosphate and sodium tetraborate, and the mobile phase B consisted of methanol: acetonitrile: water = 45:45:10 (v/v/v). The gradient elution program was (volume ratio of mobile phase B) 5% for 0–0.35 min, 5%–57% for 0.35–13.40 min, 57%–100% for 13.40–13.50 min, 100% for 13.50–15.70 min, 100–5% for 15.70–15.80 min and 5% for 15.80–18 min at 40°C with a flow rate of 1 ml min^{-1} and a detection wavelength of 338 nm.

Glucose concentration was determined using a biosensor (Su et al., 2018).

Results and discussion

Effect of phosphate concentration on L-threonine synthesis in fed-batch culture mode

From the results of the PB experiments in shake flask, phosphate is the most significant factor affecting Thr synthesis (Supplementary Table S3). In order to more truly show the effect of phosphate on Thr synthesis, effect of addition of phosphate on Thr accumulation in 5 L fermentor was investigated.

As can be seen from Figure 2A, when 4.8 g L^{-1} phosphate was added, the maximum biomass (OD_{600}) and Thr production were 56 and 42.3 g L^{-1}, respectively. With the increment of phosphate amounts up to 24.8 g L^{-1}, the biomass (OD_{600}) was improved to 76, while the change of Thr did not keep increasing (Figure 2E). The optimal concentration range of phosphate for the maximum Thr production is 9.8–14.8 g L^{-1} (Figures 2A–E), which indicates that the increased phosphate concentration may cause the migrating of C metabolic flow to the growth of biomass rather than the increase of target products. Phosphorus element is one of the core elements in microbial growth and metabolism (Anandan et al., 2014), which is involved in the composition of nucleic acids, cell membranes and high-energy phosphate compounds in life activities, and is also an important player in central metabolic pathways (Lai et al., 2012). The present study confirms that the addition of phosphate promotes the growth of cell and affects the accumulation of Thr. However, the details of phosphorus affecting metabolic flux are not very clear, which should be solved by determination of the metabolic flow distrubution.

Calculation of the metabolic flow at different phosphate concentration

Cells were cultured under fed-batch mode with different initial phosphate concentration and samples were collected at the pseudo-steady state period. Contents of glucose, Thr, L-Val, L-Tyr and L-Met were measured and the rates of metabolite consumption and accumulation were calculated as shown in Supplementary Table S4. Lingo software was used to estimate the metabolic flow distribution in the metabolic network and linear programming was performed in Excel to obtain the ideal metabolic flow distribution for Thr biosynthesis (Table 1). It can be calculated that the conversion rates of glucose to Thr are 45.4% under 9.8 g L^{-1} phosphate and 27.9% under 24.8 g L^{-1} phosphate, which are much lower than that value (73.3%) under ideal condition. With the increase of phosphate concentration, the flow rate on the branch of Thr synthesis decreases, causing the increase of by-products.

Metabolic flow analysis of key nodes in L-threonine synthesis pathway

At Glc6P nodes, when phosphate concentration changed from 9.8 g L^{-1}–24.8 g L^{-1}, the r8 flow [C/mmol·(L h)$^{-1}$] to the

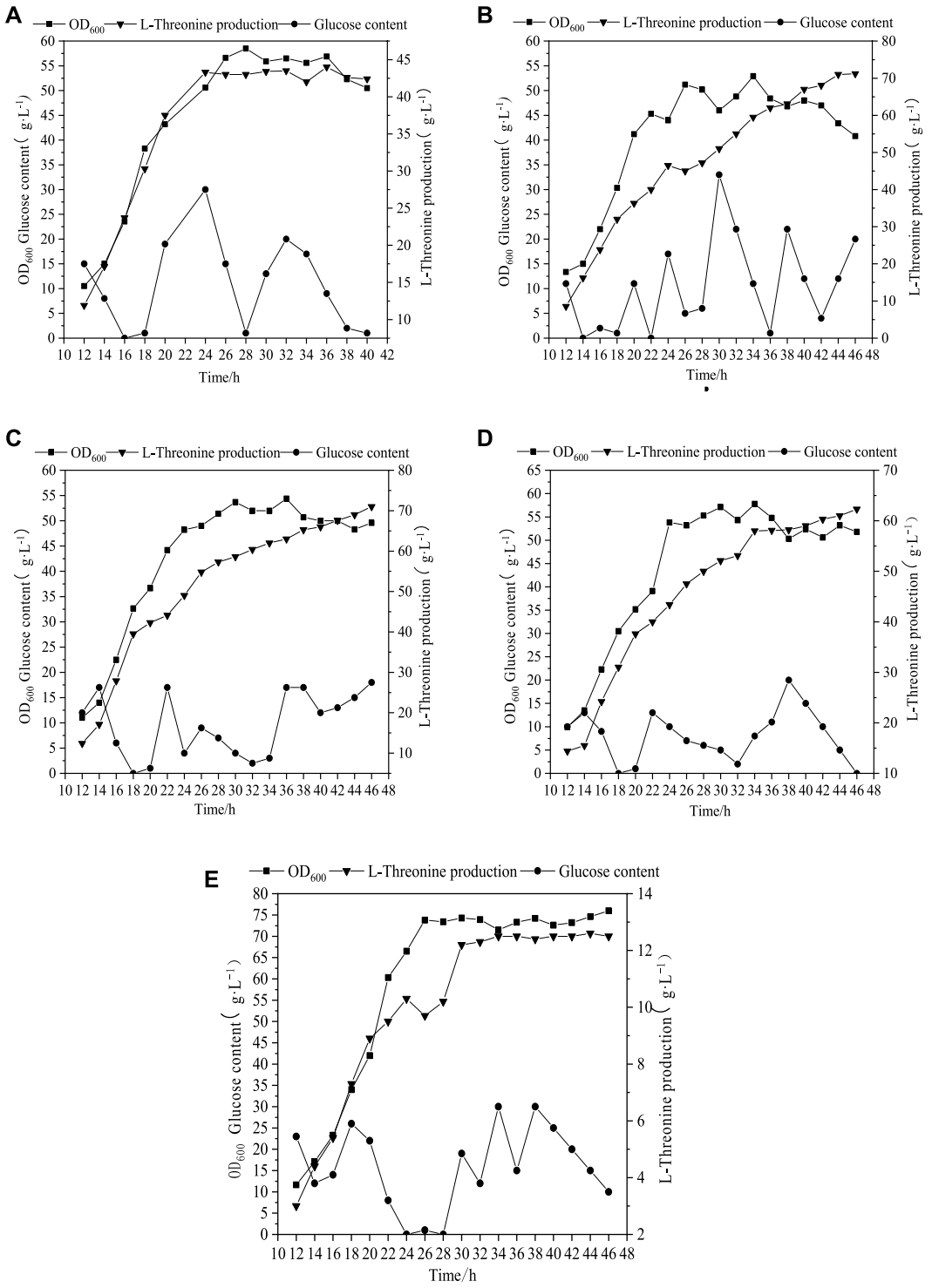


FIGURE 2
Profile of Thr production under different initial phosphate concentration. **(A)** (4.8 g L^{-1}), **(B)** (9.8 g L^{-1}), **(C)** (14.8 g L^{-1}), **(D)** (19.8 g L^{-1}), **(E)** (24.8 g L^{-1}).

HMP pathway decreased from 24.6 to 15, while the r2 flow [C/mmol·(L h)$^{-1}$] increased from 75.4 to 85 (Figure 3A). Considering the ideal metabolic flow (r2 = 0), it is inferred that enhancement of the HMP pathway metabolic flow can lead to an increase in the target metabolic flow (r26), which causes the conversion rate of glucose to Thr increased from 28% to 45%. At

TABLE 1 Metabolic flow distribution of Thr biosynthesis network.

Reaction no.	Metabolic flux distribution calculated [mmol·(L h)$^{-1}$]		Theoretical metabolic flux distribution [mmol·(L h)$^{-1}$]	Reaction no.	Metabolic flux distribution calculated [mmol·(L h)$^{-1}$]		Theoretical metabolic flux distribution [mmol·(L h)$^{-1}$]
	9.80 g L^{-1}	24.80 g L^{-1}			9.80 g L^{-1}	24.80 g L^{-1}	
r1	100.00	100.00	100.00	r14	1.08	7.083	0
r2	75.40	84.79	0	r15	1.08	7.083	0
r3	91.44	92.57	66.67	r16	0.16	9.72	0
r4	181.80	178.06	166.67	r17	0.16	9.72	0
r5	181.80	178.06	166.67	r18	46.05	44.17	73.33
r6	134.67	126.81	93.33	r19	134.52	117.08	93.33
r7	134.52	117.08	93.33	r20	134.36	107.36	93.33
r8	24.60	15.208	100	r21	134.36	107.36	93.33
r9	16.04	7.78	66.67	r22	46.05	44.17	73.33
r10	8.56	7.43	33.33	r23	45.89	34.44	73.33
r11	8.56	7.43	33.33	r24	45.89	34.44	73.33
r12	7.48	0.35	33.33	r25	0.54	6.53	0
r13	8.56	7.43	33.33	r26	45.35	27.92	73.33

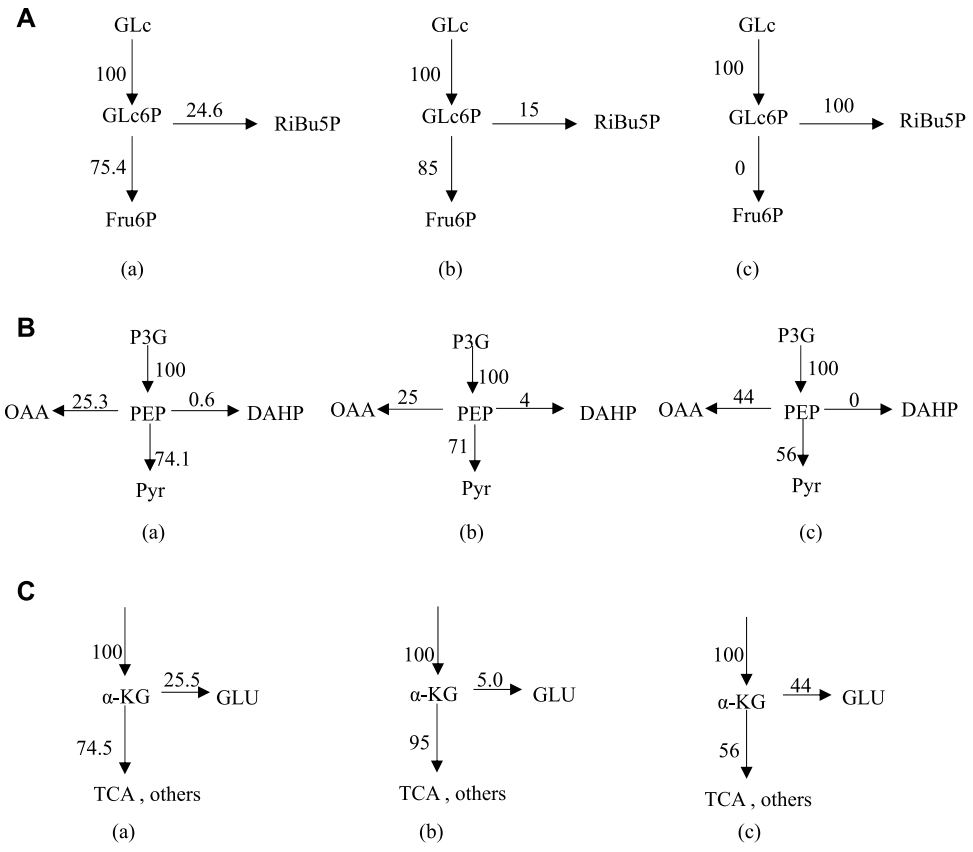


FIGURE 3
Metabolic flux distribution at Glc6P node **(A)**, PEP node **(B)** and α-KG node **(C)** under different conditions; **(A)** 9.8 g L^{-1} phosphate, **(B)** 24.8 g L^{-1} phosphate, **(C)** ideal condition. Enzyme control analysis for metabolic flux in Thr biosynthesis network.

TABLE 2 Correlation analyses between variable enzymes.

Enzyme		PEPC	HK	FBA	MDH	PK	G6PD
PEPC	Pearson correlation coefficient	1	−0.706	−0.066	−0.421	0.915[a]	0.776
	Sig. (2-tailed)		0.117	0.901	0.405	0.010	0.070
	Number	6	6	6	6	6	6
HK	Pearson correlation coefficient	−0.706	1	0.008	0.602	−0.673	−0.390
	Sig. (2-tailed)	0.117		0.989	0.206	0.143	0.445
	Number	6	6	6	6	6	6
FBA	Pearson correlation coefficient	−0.066	0.008	1	0.669	−0.133	−0.077
	Sig. (2-tailed)	0.901	0.989		0.146	0.801	0.885
	Number	6	6	6	6	6	6
MDH	Pearson correlation coefficient	−0.421	0.602	0.669	1	−0.370	0.019
	Sig. (2-tailed)	0.405	0.206	0.146		0.470	0.972
	Number	6	6	6	6	6	6
PK	Pearson correlation coefficient	0.915[a]	−0.673	−0.133	−0.370	1	0.804
	Sig. (2-tailed)	0.010	0.143	0.801	0.470		0.054
	Number	6	6	6	6	6	6
G6PD	Pearson correlation coefficient	0.776	−0.390	−0.077	0.019	0.804	1
	Sig. (2-tailed)	0.070	0.445	0.885	0.972	0.054	
	Number	6	6	6	6	6	6

[a]Denotes significance at 0.05 level (two-tailed).

PEP node (Figure 3B), 74.1% and 71% of the carbon flux (100% PEP) went to Pyr in the presence of phosphate at 9.8 g L^{-1} and 24.8 g L^{-1}, respectively. However, metabolic flow distribution under the ideal condition is 44% of the carbon flux going into OAA and 56% going to the Pyr, indicating that reduction of the flux to the TCA may help improve the synthesis of Thr. An increase in the metabolic flow of aromatic amino acid synthesis through overexpression of phosphoenolpyruvate synthase has been reported (FARMER and LIAO, 1997), thereby indirectly increasing Thr production and reducing by-product production. From Figure 3C, it can be seen that 25.5% and 5% of α-KG are catalyzed by L-glutamate dehydrogenase and L-glutamate synthase to form L-glutamate in the presence of phosphate at 9.8 g L^{-1} and 24.8 g L^{-1}, respectively. The remaining 74.5% and 95% of α-KG enters the TCA cycle under the above conditions, respectively. However, 56% metabolic flow enters the TCA cycle (r20) and 44% is used for the synthesis of L-glutamate (r22) under the ideal condition. Thus, L-glutamate is a key intermediate to provide sufficient substrate for the transamination required for the synthesis of the aspartate group of amino acids, the precursor of Thr.

Firstly, correlation analysis was performed for the six enzyme in Thr biosynthesis network and correlation coefficients were calculated for them (Table 2). The correlation coefficient between PK and PEPC is 0.915, which shows statistically significant.

To avoid the overlap of information between these enzymes, their activities of six enzymes selected in Thr metabolic network were subjected to principal component analysis (Table 3). The cumulative contribution of the first three principal components reached 95.9%, thus they are able to represent the information of all the initial indicators.

To illustrate the range of the respective original variables represented by each principal component, an analysis was performed using the principal component matrix, as shown in Table 4. PEPC, PK and G6PD show higher indices on the first principal component, indicating that information on these enzymes in the first principal component may act as descriptive indicators of Thr flux; Fructose-1,6-bisphosphate aldolase and MDH show higher indices on the second principal component, indicating that information on these enzymes in the second principal component may act as descriptive indicators of Thr flux. The higher indices shown for HK on the third principal component indicate that the information of this enzyme on the third principal component is able to serve as a descriptive indicator of Thr flux. These three linearly unrelated principal components mainly cover the full range of information and are able to replace the initial six variables.

Each score in the score coefficient matrix only represents the correlation coefficient between the principal component and the corresponding variable, which cannot represent the control effect of each enzyme on Thr flux. Under these circumstances, the coefficient matrix corresponding to each indicator in the principal component can be obtained by dividing the data in Table 5 by the corresponding eigenvalue of the principal component and opening the square root. Then the proportion

TABLE 3 Principal ingredient extraction.

Ingredients	Initial eigenvalue			Extraction of sum of squares of loads		
	Total	Percentage variance	Cumulative %	Total	Percentage variance	Cumulative %
1	3.408	56.799	56.799	3.408	56.799	56.799
2	1.549	25.824	82.623	1.549	25.824	82.623
3	0.797	13.283	95.906	1.549	13.283	95.906
4	0.159	2.653	98.559	0.159	2.653	98.559
5	0.086	1.441	100.000	0.086	1.441	100.000
6	−5.053E-16	−8.421E-15	100.000	−5.053E-16	−8.421E-15	100.000

TABLE 4 Matrix of principal components.

Enzyme	Ingredients		
	1	2	3
PEPC	0.946	0.198	−0.007
HK	−0.808	0.051	0.533
FBA	−0.266	0.815	−0.507
MDH	−0.580	0.778	0.218
PK	0.941	0.198	0.099
G6PD	0.754	0.446	0.445

TABLE 5 Matrix of principal component scoring coefficient.

Enzyme	Ingredients		
	1	2	3
PEPC	0.278	0.128	−0.008
HK	−0.237	0.033	0.669
FBA	−0.078	0.526	−0.636
MDH	−0.170	0.502	0.273
PK	0.276	0.128	0.125
G6PD	0.221	0.288	0.558

of the eigenvalue corresponding to each principal component to the sum of the total eigenvalues of the extracted principal components was used as the weight to calculate the comprehensive expression coefficient of each principal component, which integrates the control effect of each enzyme on Thr flux. Their values were normalized and recorded as the control coefficient (CCP), as shown in Table 6. G6PD played the most dominant role in controlling Thr flux, while fructose-1,6-bisphosphate aldolase negatively regulated Thr flux. In conclusion, the control coefficients constructed based on quantitative genetic methods reflects the metabolic regulation of Thr flux by the six enzymes.

Metabolomic analysis of L-threonine synthesis pathway

Differential metabolite analysis

With the detection of GC-MS, the metabolites were determined by qualitative and quantitative analyses (Table 7). Organic acids, amino acids, sugars and alcohols represent the main part of metabolites on the aspect of contents. An unreplicated two-way ANOVA was performed to determine whether significant differences exist between strains TWF001 and W3110 in terms of the contents of each metabolite.

Principal component analysis was carried out for the experimental strain (sample T) and the control strain (sample W) and Figure 4A shows that the differences between the two groups of metabolites are significantly obvious. The principal component interpretation rate $R^2(X) = 0.868 > 0.5$ indicates a good model fit, and $Q^2 = 0.5651$, which is less different from $R^2(X)$, indicating the stability of the fitted equations. PCA is an unsupervised analysis method that cannot ignore the errors of each group and eliminate random errors, so it is necessary to adopt a supervised approach to identify differences between groups. Orthogonal partial least squares discriminant analysis (OPLS-DA) is a supervised analysis method that combines partial least squares and discriminant analysis. OPLS-DA classifies the sample variable matrix X as both correlated and uncorrelated with Y, and removes the irrelevant variation variables, which enables a comprehensive analysis of between- and within-group variance. From Figure 4B, here $R^2(X)$ and $R^2(Y)$ describe the explanatory rate of the model, and $Q^2 = 0.899$ represents the predictive power of the model. These three indicators are close to 1, indicating a good reliability of this model (Li et al., 2021). The samples in the Figure 4B are all within the 95% confidence interval and the two class of samples are significantly differentiated with the dispersion in the T sample being greater than the dispersion in the W sample.

TABLE 6 Comprehensive express coefficient matrix and control coefficients based on principal component analysis.

Enzyme	Principal component expression coefficient matrix			Principal component control expression coefficients	CCP
	A1	A2	A3		
PEPC	0.151	0.103	−0.009	0.116	0.058
HK	−0.128	0.265	0.749	0.1	0.05
FBA	−0.0423	0.422	−0.712	−0.01045	−0.005
MDH	−0.0921	0.403	0.306	0.096	0.048
PK	0.15	0.103	0.140	0.136	0.068
G6DH	0.12	0.231	0.625	0.22	0.11

TABLE 7 Metabolite distribution in experimental and control strains.

Retention time		12 h		24 h		36 h	
(min)	metabolites	W3110	TWF001	W3110	TWF001	W3110	TWF001
7.79	Carbamic acid	0.0423 ± 0.016	0.0843 ± 0.015	0.0668 ± 0.017	0.0468 ± 0.005[a]	0.0667 ± 0.027	0.0542 ± 0.005[b]
8.02	Lactic acid	0.3162 ± 0.027	1.8662 ± 0.26[b]	5.3 ± 0.47	4.1 ± 0.66	4.4571 ± 0.33	2.1569 ± 0.47
8.31	Glycolic acid	0.0061 ± 0.00029	0.0114 ± 0.00304[b]	0.0534 ± 0.007	0.0357 ± 0.006	0.061 ± 0.0037	0.0474 ± 0.01[a]
8.49	Valine	0.0454 ± 0.01764	0.0468 ± 0.024	0.0218 ± 0.015	0.0826 ± 0.031	0.0372 ± 0.1302	0.0329 ± 0.011
8.96	Alanine	0.0477 ± 0.01	0.0527 ± 0.01	0.1164 ± 0.033	0.0203 ± 0.012	0.124 ± 0.03	0.03449 ± 0.005[b]
9.75	Acetic acid	0.0876 ± 0.008	0.1356 ± 0.02	0.0628 ± 0.011	0.1007 ± 0.016	0.0752 ± 0.0025	0.0794 ± 0.017
10.49	Pentanoic acid	0.004 ± 0.0006	0.0688 ± 0.014[a]	0.0189 ± 0.0027	0.081 ± 0.011[b]	0.0144 ± 0.0028	0.0340 ± 0.006
10.65	Leucine	0.0164 ± 0.015	0.0148 ± 0.022	0.012 ± 0.002	0.0095 ± 0.007	0.0127 ± 0.0022	0.0945 ± 0.016[b]
11.42	Ciniic acid	0.0047 ± 0.00056	0.0344 ± 0.00714[a]	0.0073 ± 0.0018	0.0804 ± 0.0272[a]	-	0.0477 ± 0.006
13.75	Thr	0.1033 ± 0.038	5.73 ± 1.16[a]	0.1058 ± 0.046	14.64 ± 1.03[a]	0.0557 ± 0.05	12.1 ± 1.48[b]
13.98	Succinic acid	0.0238 ± 0.00247	0.8841 ± 0.07[b]	0.9643 ± 0.098	1.4598 ± 0.128	0.9193 ± 0.035	0.7538 ± 0.1129[a]
14.49	Glycolic acid	0.0188 ± 0.00654	0.1003 ± 0.00472	0.2145 ± 0.036	0.2412 ± 0.02[a]	0.2056 ± 0.03	0.1332 ± 0.02
14.93	Homoserine	-	0.0913 ± 0.005	-	0.2784 ± 0.05	-	0.2620 ± 0.065
15.40	Malic acid	0.5801 ± 0.05	0.0067 ± 0.001[a]	0.0362 ± 0.0078	0.2784 ± 0.05	0.0528 ± 0.0025[a]	0.0399 ± 0.014
17.52	Citric acid	0.0136 ± 0.0037	0.0248 ± 0.002	0.1563 ± 0.01	0.0383 ± 0.007	0.1640 ± 0.01	0.0322 ± 0.006[a]
18.20	L-Glutamic acid	0.0475 ± 0.0056	0.0129 ± 0.00412	0.8073 ± 0.103	0.0512 ± 0.007[b]	1.5650 ± 0.52	0.0753 ± 0.016[a]
18.44	Sucralose alcohol	0.0067 ± 0.00028	0.0243 ± 0.037[a]	-	0.0168 ± 0.003	-	0.0234 ± 0.002
21.57	Arabinose	0.1378 ± 0.201	0.0556 ± 0.043	0.1337 ± 0.14	0.0717 ± 0.068	0.1197 ± 0.03[a]	0.0437 ± 0.004[b]
24.48	Tartar	0.0281 ± 0.0035	0.0052 ± 0.001	0.0472 ± 0.011	0.0237 ± 0.002[b]	0.0418 ± 0.0042	0.0315 ± 0.011
27.60	Palmitic acid	0.4538 ± 0.15	0.3034 ± 0.09	0.4939 ± 0.127	0.1728 ± 0.08	0.4113 ± 0.188	0.3219 ± 0.114
27.79	Aspartic acid	0.1493 ± 0.02	0.0029 ± 0.00026[a]	0.1425 ± 0.12	0.0152 ± 0.005[b]	0.2632 ± 0.05	0.0201 ± 0.0022[b]
28.82	Inositol	0.0152 ± 0.00072	0.0088 ± 0.0005	0.0347 ± 0.007	0.0173 ± 0.006	0.0237 ± 0.003	0.0177 ± 0.0032
30.56	Stearic acid	0.2889 ± 0.12	0.1933 ± 0.06	0.2829 ± 0.15	0.1345 ± 0.039[b]	0.2882 ± 0.143	0.2281 ± 0.03[a]
32.80	Mannitol	0.0124 ± 0.0008	0.56 ± 0.11[b]	0.045 ± 0.007	0.148 ± 0.013	0.3133 ± 0.017	0.3995 ± 0.026
35.37	Monopalmitin	0.2102 ± 0.039	0.2334 ± 0.063	0.2153 ± 0.11	0.1428 ± 0.07	0.3178 ± 0.1	0.1639 ± 0.1062
37.72	Glyceryl monostearate	0.1450 ± 0.018	0.1636 ± 0.022	0.1193 ± 0.03	0.1281 ± 0.03	0.5383 ± 0.063	0.1409 ± 0.07

Note: Metabolite contents are g L^{-1}.
[a]Refers to $p < 0.05$.
[b]Refers to $p < 0.01$.

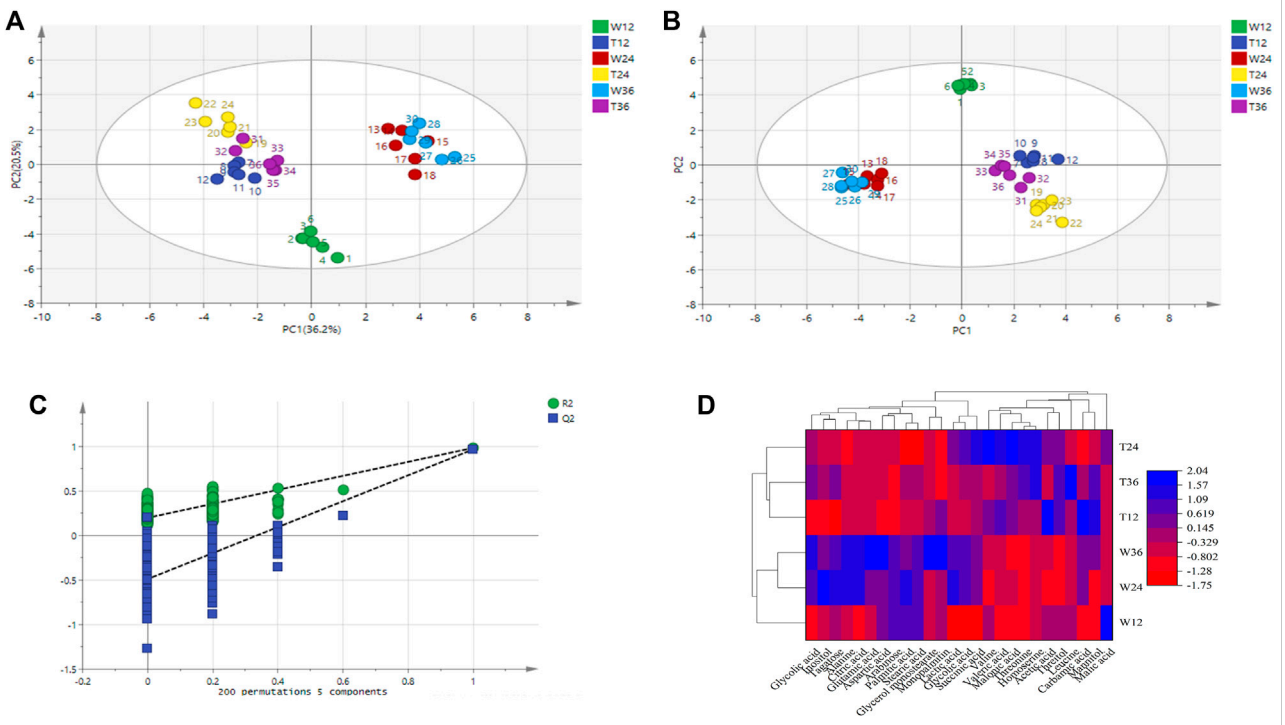


FIGURE 4
Metabolite difference in strain TWF001 and W3110. **(A)** Score plot of metabolite principal component analysis. **(B)** OPLS-DA score plot of metabolite principal component analysis. **(C)** OPLS-DA displacement test plot. **(D)** Cluster analysis of the two strains. Red represents the zone of low concentration of differential metabolites and blue represents the zone of high concentration of differential metabolites.

An OPLS-DA model with good predictability and fit has been developed, while a 200-response permutation test model was also necessary to be constructed to prevent the model from over-fitting (Figure 4C). Here R^2(Y) (interpretability of the Y variable) and Q^2 (model predictability) are important parameters for model evaluation, and their regression lines are crossed with the horizontal coordinates or less than 0, which indicates that the model is relatively accurate. The R^2(Y) value (0.94) in the OPLS-DA model is approximately equal to 1, indicating that the model is relatively reliable and is able to reflect the real situation of the sample data. The intercept of Q^2 on the *Y*-axis is negative, and the R^2 = 0.218 and Q^2 = -0.483 obtained from the replacement test are smaller than the initial values of R^2 and Q^2 in the OPLS-DA model, which suggests that the model is not over-fitted.

The data set was scaled by the heatmap package in Origin 2018 and a cluster analysis was made based on the differential metabolites and their contents to further reflect the metabolic differences between the two strains. Figure 4D shows that the same strain is more correlated even at different culture times, while the metabolic differences between TWF001 and W3110 are high, which is consistent with the results of the principal component analysis.

Difference analysis of L-threonine metabolic network and potential biomarkers

Figure 5A shows that the main metabolites contents that differed between TWF001 and W3110 include amino acids, organic acids, fatty acids, sugars, alcohols and other substances. The contents of L-alanine, L-glutamic acid, L-aspartic acid and fatty acid compounds were significantly lower in TWF001 when compared with the control strain. TWF001 produced higher content of Thr with less carbon metabolic flow to other by-product amino acids. However, L-glutamate plays an important role in the pathway of L-aspartate, the precursor of Thr synthesis, thus up-regulation of the relative genes of glutamate synthesis has several sets of effects on Thr synthesis (Zhao et al., 2020). As a by-product, the increased concentration of L-leucine is detrimental to Thr synthesis, so Lee (Lee et al., 2007)et al. weakened the enzyme activity by site directed mutation of *ilvA* (encoding Thr dehydratase), and thereby reducing L-isoleucine and L-leucine concentrations to prevent Thr degradation. In addition, the higher C flow to by-product amino acids and fatty acid branched pathways in W3110 eventually led to lower Thr level in that strain.

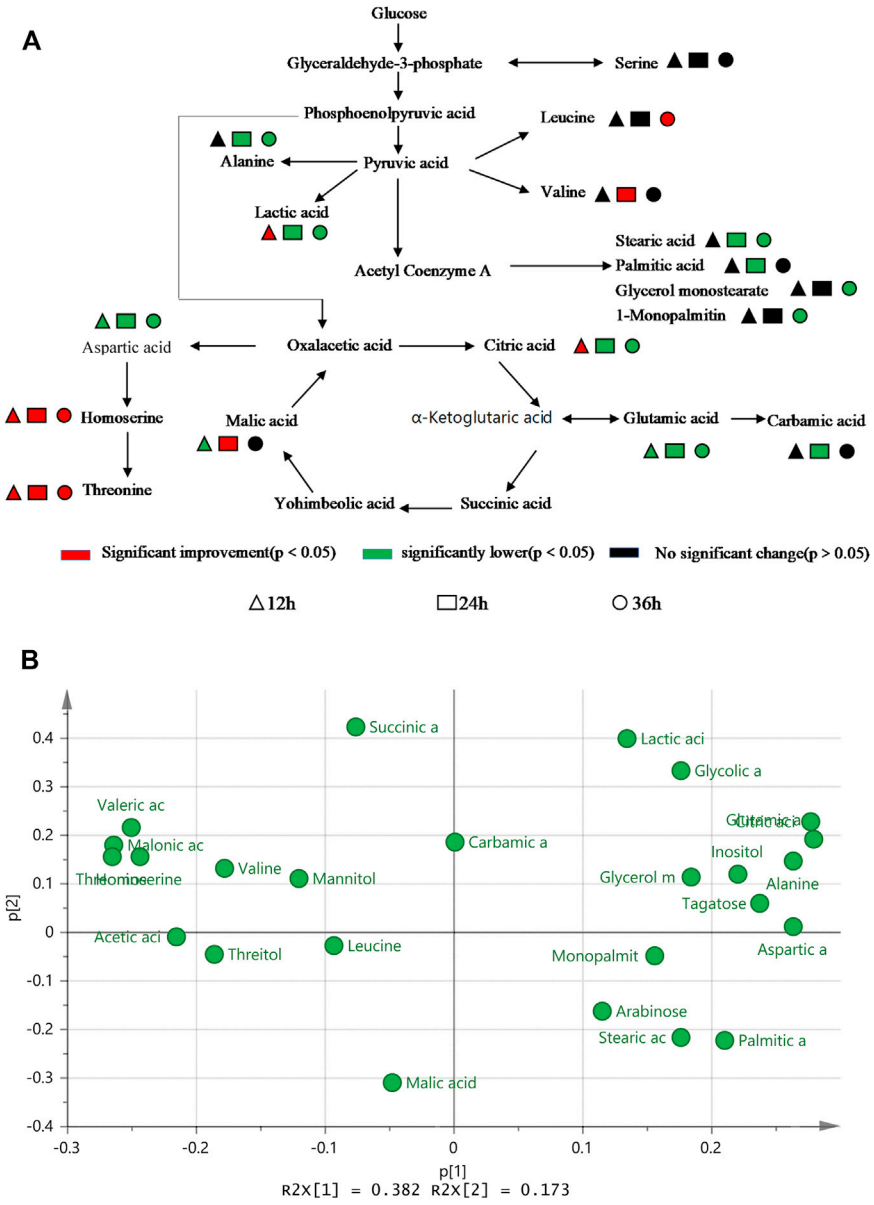


FIGURE 5
(A) Difference in metabolic level between TWF001 and W3110. **(B)** Load plot for metabolite principal component analysis between TWF001 and W3110.

When the intracellular environment is disturbed by the external factors, the metabolic change will occur to adapt to this disturbance and a new metabolic equilibrium is achieved. In PCA analysis, the load diagram represents the magnitude of the contribution made to differentiate between experimental and control groups, with the further away from the origin simply indicating that the metabolite is making a more significant contribution and can be regarded as a biomarker. As shown in Figure 5B, citric acid, L-glutamic acid and inositol are the furthest from the origin, and the levels of these metabolites are proportional to the synthesis of Thr within a certain range. Therefore, these metabolites can be good biomarkers to be added in the medium to regulate the synthesis of Thr.

A case of metabolic regulation based on metabolic analysis of Thr biosynthesis pathway

From above analyses, some key nods, enzymes and biomarkers have been identified as important targets for the production of Thr. Here, we selected addition of L-glutamic acid as an example to verify

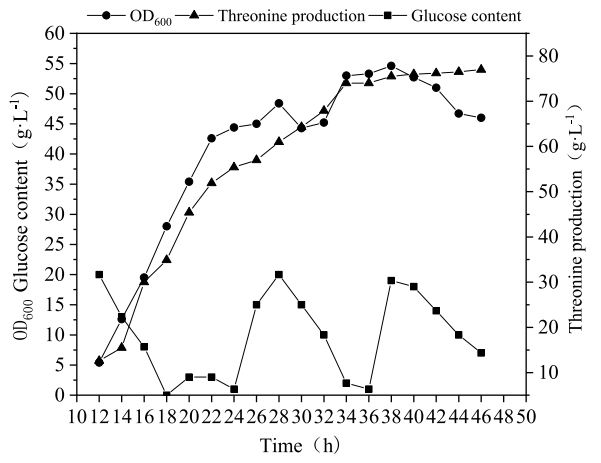


FIGURE 6
| Profile of Thr production under fed-batch mode with the addition of 8 L-glutamic acid

the reliability of the biosynthesis pathway analyses. L-glutamic acid was added into medium and fed-batch fermentation was conducted to evaluate the performance. As can be seen from Figure 6, the highest OD_{600} value of was not improved, but the Thr production was promoted up to 77 g L^{-1}, 10% higher than that without addition of L-glutamic acid. This fact indicates that L-glutamic acid may function by migrating the metabolic flow to Thr rather than enhancing biomass.

Some metabolic regulators, such as glycine, sodium nitrate and sulphate, are activators or inhibitors of key enzymes in the pathway. However, addition of them is often not sufficient to achieve the desired effect. In addition to environmental disturbance, some key enzymes need to be engineered to change the flow of metabolic flux in order to achieve the most reasonable flux distribution for Thr synthesis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

QY, WL, SC and HC conceived and planned the experiments. QY and WL performed the conceptualization, methodology, experiments, data analyses, and writing. WL, DC and PY reviewed and edited the draft. All authors approved the final manuscript.

Funding

The authors are grateful for the financial support provided by the Key Technology Research Plan Project of the Inner Mongolia Autonomous Region (2019GG302), the Open Project Funding of State Key Laboratory of Biocatalysis and Enzyme Engineering (SKLBEE2021003) and the Program of the Key Laboratory of Industrial Biotechnology, the Ministry of Education, China (no. KLIB-KF202001).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.1010931/full#supplementary-material

References

Anandan, P., Parthipan, G., Pazhanivel, K., Ravi, G., and Jayavel, R. (2014). Growth and characterization of potassium halides mixed l-arginine phosphate monohydrate semi organic nonlinear optical single crystals. *Optik* 125 (1), 8–10. doi:10.1016/j.ijleo.2013.05.170

Antoniewicz, M. R. (2021). A guide to metabolic flux analysis in metabolic engineering: Methods, tools and applications. *Metab. Eng.* 63, 2–12. doi:10.1016/j.ymben.2020.11.002

Debabov, V. G., Kozlov, J. I., Zhdanova, N. I., Khurges, E. M., Yankovsky, N. K., Rozinov, M. N., et al. (1981). *Method for preparing strains which produce aminoacids*. New York: United States Patent.

Dong, X., Quinn Peter, J., and Wang, X. (2011). Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for the production of L-threonine. *Biotechnol. Adv.* 29 (1), 11–23. doi:10.1016/j.biotechadv.2010.07.009

Farmer, W. R., and Liao, J. A. M. E. S. C. (1997). Reduction of aerobic acetate production by *Escherichia coli* [J]. *Appl. Environ. Microbiol.* 63 (8), 3205–3210. doi:10.1128/aem.63.8.3205-3210.1997

Follettie, M., Shin, H., and Sinskey, A. (1988). Organization and regulation of the *Corynebacterium glutamicum* hom-thrB and thrC loci [J]. *Mol. Microbiol.* 2 (1), 53–62. doi:10.1111/j.1365-2958.1988.tb00006.x

Jonsson, P., Gullberg, J., Nordström, A., Kusano, M., Kowalczyk, M., Sjostrom, M., et al. (2004). A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. *Anal. Chem.* 76 (6), 1738–1745. doi:10.1021/ac0352427

Koek, M. M., Bakels, F., Engel, W., van den Maagdenberg, A., Ferrari, M. D., Coulier, L., et al. (2010). Metabolic profiling of ultrasmall sample volumes with GC/MS: From microliter to nanoliter samples. *Anal. Chem.* 82 (1), 156–162. doi:10.1021/ac9015787

Komatsubara, S., Murata, K., Kisumi, M., and Chibata, I. (1978). Threonine degradation by *Serratia marcescens*. *J. Bacteriol.* 135 (2), 318–323. doi:10.1128/jb.135.2.318-323.1978

Lai, S., Zhang, Y., Liu, S., Liang, Y., Shang, X., Chai, X., et al. (2012). Metabolic engineering and flux analysis of *Corynebacterium glutamicum* for L-serine production [J]. *Sci. China Life Sci.* 55 (4), 283–290. doi:10.1007/s11427-012-4304-0

Lee, J. H., Lee, D. E., Lee, B. U., and Kim, H. S. (2003). Global analyses of transcriptomes and proteomes of a parent strain and an l -Threonine-Overproducing mutant strain. *J. Bacteriol.* 185 (18), 5442–5451. doi:10.1128/jb.185.18.5442-5451.2003

Lee, J. H., Sung, B., Kim, M., Blattner, F. R., Yoon, B. H., Kim, J. H., et al. (2009). Metabolic engineering of a reduced-genome strain of *Escherichia coli* for L-threonine production. *Microb. Cell. Factories* 8 (1), 1–12. doi:10.1186/1475-2859-8-2

Lee, K. H., Park, J. H., Kim, T. Y., and Lee, S. Y. (2007). Systems metabolic engineering of *Escherichia coli* for L-threonine production [J]. *Mol. Syst. Biol.* 3 (1), 149. doi:10.1038/msb4100196

Leuchtenberger, W., Huthmacher, K., and Drauz, K. (2005). Biotechnological production of amino acids and derivatives: Current status and prospects. *Appl. Microbiol. Biotechnol.* 69 (1), 1–8. doi:10.1007/s00253-005-0155-y

Li, Z., 2, Zhao, L., Xie, F., Yang, C., Jayamanne, V. S., Tan, H., et al. (2021). Study of assessment of green tea' grades in GC-MS determination of aromatic components based on principal component analysis (PCA) [J]. *J. Food Process. Preserv.* 45 (1), e15047. doi:10.1111/jfpp.15047

Liu, H., Wang, Y., Hou, Y., and Li, Z. (2020). Fitness of chassis cells and metabolic pathways for l-cysteine overproduction in *Escherichia coli* [J]. *J. Agric. Food Chem.* 68 (50), 14928–14937. doi:10.1021/acs.jafc.0c06134

Luo, W., Gong, Z., Li, N., Zhao, Y., Zhang, H., Yang, X., et al. (2020). A negative regulator of carotenogenesis in *Blakeslea trispora* [J]. *Appl. Environ. Microbiol.* 86 (6), e02462–02419.

Luo, W., Hu, J., Lu, J., Zhang, H., Wang, X., Liu, Y., et al. (2021). One pot cascade synthesis of L-2-aminobutyric acid employing ω-transaminase from *Paracoccus pantotrophus* [J]. *Mol. Catal.* 515, 111890. doi:10.1016/j.mcat.2021.111890

OlafKrömera, J., Wittmanna, C., Schröderb, H., and Heinzle, E. (2006). Metabolic pathway analysis for rational design of L-methionine production by *Escherichia coli* and Corynebacterium glutamicum[J]. *Metab. Eng.* 8 (4), 353–369. doi:10.1016/j.ymben.2006.02.001

Park, C. H., Park, Y. E., Yeo, H. J., Yoon, J. S., Park, S. Y., Kim, J. K., et al. (2021). Comparative analysis of secondary metabolites and metabolic profiling between diploid and tetraploid *Morus alba* L. *J. Agric. Food Chem.* 69 (4), 1300–1307. doi:10.1021/acs.jafc.0c06863

Park, J. H, Y. L. S., and Lee, S. Y. (2010). Metabolic pathways and fermentative production of L-aspartate family amino acids. *Biotechnol. J.* 5 (6), 560–577. doi:10.1002/biot.201000032

Picone, G., Mezzetti, B., Babini, E., Capocasa, F., Placucci, G., and Capozzi, F. (2011). Unsupervised principal component analysis of NMR metabolic profiles for the assessment of substantial equivalence of transgenic grapes (Vitis vinifera). *J. Agric. Food Chem.* 59 (17), 9271–9279. doi:10.1021/jf2020717

Schmid, J. W., Mauch, K., Reuss, M., Gilles, E. D., and Kremling, A. (2004). Metabolic design based on a coupled gene expression—Metabolic network model of tryptophan production in *Escherichia coli* [J]. *Metab. Eng.* 6 (4), 364–377. doi:10.1016/j.ymben.2004.06.003

Song, K. H., Lee, H. H., and Hyun, H. H. (2000). Characterization of salt-tolerant mutant for enhancement of L-threonine production in *Escherichia coli* [J]. *Appl. Microbiol. Biotechnol.* 54 (5), 647–651. doi:10.1007/s002530000417

Su, Y., Guo, Q. Q, S. W., Wang, S., Zhang, X., and Wang, J. (2018). Effects of betaine supplementation on L-threonine fed-batch fermentation by *Escherichia coli* [J]. *Bioprocess Biosyst. Eng.* 41 (10), 1509–1518. doi:10.1007/s00449-018-1978-0

Toya, Y., Kono, N., Arakawa, K., and Tomita, M. (2011). Metabolic flux analysis and visualization. *J. Proteome Res.* 10 (8), 3313–3323. doi:10.1021/pr2002885

Vo, T-C., Lee, J., Son, B-G., Kim, C. K., Kim, H., and Lim, K. B. (2018). Phenotypic correlation analysis of Phalaenopsis reciprocal F1 hybrids using SPSS and principal component analysis (PCA). *Acta Hortic.* 1262, 213–218. doi:10.17660/actahortic.2019.1262.28

Wang, J., Cheng, L-K., and Chen, N. (2014). High-level production of L-threonine by recombinant *Escherichia coli* with combined feeding strategies [J]. *Biotechnol. Biotechnol. Equip.* 28 (3), 495–501. doi:10.1080/13102818.2014.927682

Yang, K., Duley, M. L., and Zhu, J. (2018). Metabolomics study reveals enhanced inhibition and metabolic dysregulation in *Escherichia coli* induced by lactobacillus acidophilus-fermented black tea extract [J]. *J. Agric. Food Chem.* 66 (6), 1386–1393. doi:10.1021/acs.jafc.7b04752

Zhao, H., Fang, Y., Wang, X., Zhao, L., Wang, J., and Li, Y. (2018). Increasing l-threonine production in *Escherichia coli* by engineering the glyoxylate shunt and the L-threonine biosynthesis pathway [J]. *Appl. Microbiol. Biotechnol.* 102 (13), 5505–5518. doi:10.1007/s00253-018-9024-3

Zhao, L., Zhang, H., Wang, X., Han, G., Ma, W., Hu, X., et al. (2020). Transcriptomic analysis of an l -threonine-producing *Escherichia coli* TWF001 [J]. *Biotechnol. Appl. Biochem.* 67 (3), 414–429. doi:10.1002/bab.1890

Zheng, H-Y., Zhang, M., Yang, C-X., Zhang, N., Wang, X. C., Yang, X. P., et al. (2020). Elevated exhaustion levels and reduced functional diversity of T cells in peripheral blood may predict severe progression in COVID-19 patients. *Cell. Mol. Immunol.* 17 (5), 541–543. doi:10.1038/s41423-020-0401-3

Zhu, W. Y., Niu, K., Liu, P., Fan, Y., Liu, Z., and Zheng, Y. (2021). Enhanced O-succinyl-l-homoserine production by recombinant *Escherichia coli* ΔIJBB* TrcmetL/pTrc-metAfbr-Trc-thrAfbr-yjeH via multilevel fermentation optimization [J]. *J. Appl. Microbiol.* 130 (6), 1960–1971. doi:10.1111/jam.14884

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 20 October 2022
DOI 10.3389/fbioe.2022.1050808

OPEN ACCESS

EDITED BY
Zhenlin Han,
University of Hawaii at Manoa,
United States

REVIEWED BY
Riyi Xu,
Guangdong Academy of Sciences,
China
Wen-Yong Lou,
South China University of Technology,
China
Shizhen Wang,
Xiamen University, China

*CORRESPONDENCE
Li-Hai Fan,
fanlh@fzu.edu.cn
Hui-Dong Zheng,
youngman@fzu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and Biotechnology

RECEIVED 22 September 2022
ACCEPTED 11 October 2022
PUBLISHED 20 October 2022

CITATION
Zheng L-J, Guo Q, Zhang Y-X, Liu C-Y, Fan L-H and Zheng H-D (2022), Engineering of Escherichia coli for D-allose fermentative synthesis from D-glucose through izumoring cascade epimerization.
Front. Bioeng. Biotechnol. 10:1050808.
doi: 10.3389/fbioe.2022.1050808

Engineering of *Escherichia coli* for D-allose fermentative synthesis from D-glucose through izumoring cascade epimerization

Ling-Jie Zheng[1,2†], Qiang Guo[1†], Ya-Xing Zhang[1], Chen-Yang Liu[1], Li-Hai Fan[1,2*] and Hui-Dong Zheng[1,2*]

[1]College of Chemical Engineering, Fujian Engineering Research Center of Advanced Manufacturing Technology for Fine Chemicals, Fuzhou University, Fuzhou, China, [2]Qingyuan Innovation Laboratory, Quanzhou, China

D-Allose is a potential alternative to sucrose in the food industries and a useful additive for the healthcare products in the future. At present, the methods for large-scale production of D-allose are still under investigation, most of which are based on *in vitro* enzyme-catalyzed Izumoring epimerization. In contrast, fermentative synthesis of D-allose has never been reported, probably due to the absence of available natural microorganisms. In this work, we co-expressed D-galactose: H^+ symporter (GalP), D-glucose isomerase (DGI), D-allulose 3-epimerase (DAE), and ribose-5-phosphate isomerase (RPI) in *Escherichia coli*, thereby constructing an *in vivo* Izumoring pathway for yielding D-allose from D-glucose. The carbon fluxes and carbon catabolite repression (CCR) were rationally regulated by knockout of FruA, PtsG, Glk, Mak, PfkA, and PfkB involved in the pathways capable of phosphorylating D-fructose, D-glucose, and fructose-6-phosphate. Moreover, the native D-allose transporter was damaged by inactivation of AlsB, thus driving the reversible Izumoring reactions towards the target product. Fermentation was performed in the M9 medium supplemented with glycerol as a carbon source and D-glucose as a substrate. The results show that the engineered *E. coli* cell factory was able to produce approximately 127.35 mg/L of D-allose after 84 h. Our achievements in the fermentative production of D-allose in this work may further promote the green manufacturing of rare sugars.

KEYWORDS

metabolic engineering, fermentation, rare sugar, biomanufacturing, cell factory

Introduction

D-Allose is a rare hexose sugar with ultra-low energy and 80% sucrose sweetness (Iga et al., 2010; Mooradian et al., 2017). Many researches have reported that D-allose shows certain efficacy in inhibiting cancers, such as ovarian cancer, hepatocellular carcinoma, pancreatic cancer, prostate cancer, and cervical cancer (Naha et al., 2008; Yamaguchi et al., 2008; Yokohira et al., 2008). Moreover, it can also act as an anti-inflammatory agent to alleviate cisplatin-induced nephrotoxicity (Miyawaki et al., 2012), as an antioxidant to prevent oxidative damage caused by reactive oxygen species (ROS) (Sun et al., 2006; Ishihara et al., 2011; Nakamura et al., 2011), and as an immunosuppressant in cryoprotection of biological cells and tissues. These beneficial physiological properties make D-allose a potential sweetener in food and healthcare products, and the production of D-allose has been gradually becoming a research focus (Lim and Oh, 2011).

The routes for D-allose synthesis are either chemical or biological, of which chemical process normally generates a variety of by-products, and the harmless disposal of the resulting waste also remains a challenge (Wang et al., 2020; Morimoto et al., 2022; Vigo et al., 2022). In contrast, biosynthesis has the advantages of high specificity, mild conditions, and environmental friendliness, and is in line with the concept of green manufacturing. Currently, most of the studies on D-allose bioproduction are based on the Izumoring enzymatic cascade (Granström et al., 2004; Izumori, 2006). Izumoring was proposed in 2004 as a strategy to synthesize rare sugars through enzyme-catalyzed epimerization between monosaccharides (Granström et al., 2004). Although reversible epimerization leads to the low conversion of substrates, this method is still the best choice for preparation of rare sugars, and has been successfully applied in the industrial production of D-allulose, a stereoisomer of D-fructose (Zhu et al., 2012; Su et al., 2018). However, it is interesting that Izumoring synthesis of rare sugars is basically achieved through *in vitro* biocatalysis rather than cell factories (Kim et al., 2006; Zhu et al., 2012). The reason may be that the substrate for Izumoring is usually D-fructose or D-glucose, which can be efficiently phosphorylated by cells, thus being allowed to enter the central metabolic pathways as a carbon source for growth (Guggisberg et al., 2014; Luo et al., 2014). Therefore, although fermentation has more cost-reducing potential than enzymatic catalysis for D-allose production, metabolic engineering of cell factories is quite important, especially the phosphorylation pathways of substrate and intermediates should be rationally reprogrammed.

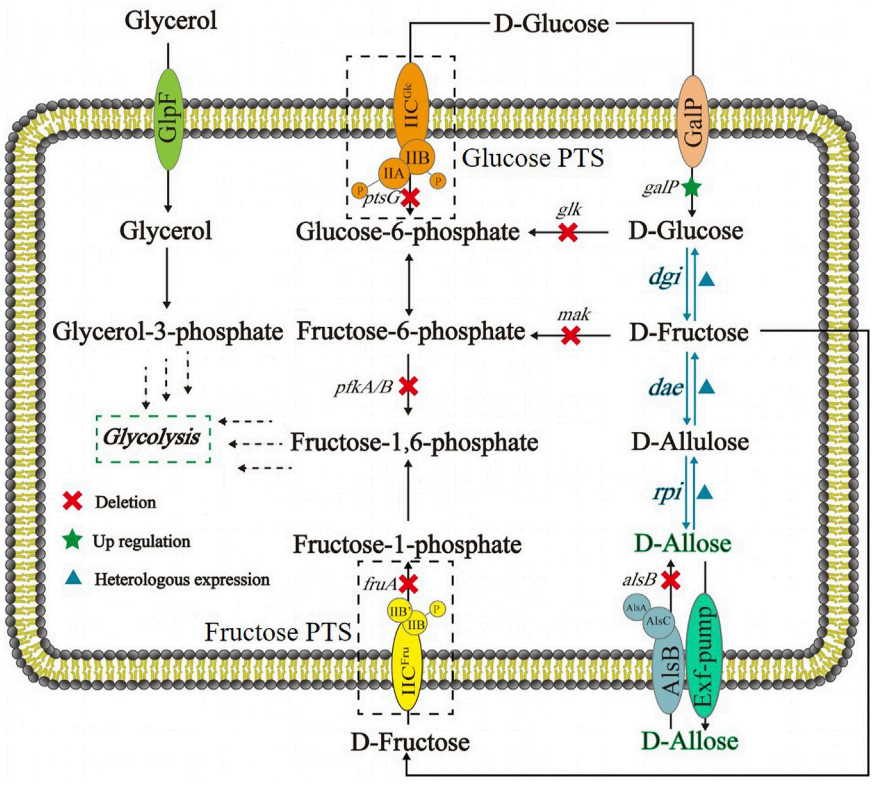


FIGURE 1
Schematic of *E. coli* cell factory for synthesizing D-allose from D-glucose based on Izumoring cascade epimerization. DGI: D-glucose isomerase; DAE: D-allulose 3-epimerase; RPI: ribose-5-phosphate isomerase; GalP: D-galactose H^+ symporter; Glucose PTS: the phosphoenolpyruvate transferase system of D-glucose; Fructose PTS: the phosphoenolpyruvate transferase system of D-fructose.

TABLE 1 Strains and plasmids used in this study.

Name	Relevant characteristics	References
Strains		
E. coli JM109 (DE3)	Wild type *E. coli*	LMAI Bio
E. coli (Empty A)	*E. coli* JM109 (DE3) harboring pETDuet-1	this study
E. coli (Empty B)	*E. coli* JM109 (DE3) harboring pETDuet-1 and pRSFDuet-1	this study
E. coli (DGI)	*E. coli* JM109 (DE3) harboring pETDuet-*dgi*	this study
E. coli (DAE)	*E. coli* JM109 (DE3) harboring pETDuet-*dae*	this study
E. coli (RPI)	*E. coli* JM109 (DE3) harboring pETDuet-*rpi*	this study
E. coli (DGI, DAE, RPI, GalP)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi* and pRSFDuet-*dae-rpi*	this study
E. coli (DGI, GalP)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi*	this study
E. coli (DGI, GalP, ΔFruA)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi* with *fruA* deleted	this study
E. coli (DGI, GalP, ΔFruA, ΔPtsG)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi* with *fruA* and *ptsG* deleted	this study
E. coli (DGI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, ΔPfkB)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi* with *fruA*, *ptsG*, *glk*, *mak*, *pfkA*, and *pfkB* deleted	this study
E. coli (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, ΔPfkB)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi* and pRSFDuet-*dae-rpi* with *fruA*, *ptsG*, *glk*, *mak*, *pfkA*, and *pfkB* deleted	this study
E. coli (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, ΔPfkB, ΔAlsB)	*E. coli* JM109 (DE3) harboring pETDuet-*galP-dgi* and pRSFDuet-*dae-rpi* with *fruA*, *ptsG*, *glk*, *mak*, *pfkA*, *pfkB*, and *alsB* deleted	this study
Plasmids		
pETDuet-1	T7 promoter, ampicillin resistance	Novagen
pRSFDuet-1	T7 promoter, kanamycin resistance	Novagen
pETDuet-*dgi*	pETDuet-1 expressing *dgi*	this study
pETDuet-*dae*	pETDuet-1 expressing *dae*	this study
pETDuet-*rpi*	pETDuet-1 expressing *rpi*	this study
pETDuet-*galP-dgi*	pETDuet-1 co-expressing *galP* and *dgi*	this study
pRSFDuet-*dae-rpi*	pRSFDuet-1 co-expressing *dae* and *rpi*	this study
pKD46, pCP20, pKD13	λ red recombination system	Guo et al. (2021)

Here we designed and constructed an *Escherichia coli* cell factory capable of synthesizing D-allose from D-glucose (Figure 1), thereby demonstrating that it is possible to produce rare sugars by use of *in vivo* Izumoring cascade epimerization (Lim and Oh, 2011; Chen et al., 2018). Briefly, to ensure that D-glucose was available in cells as a substrate in its unphosphorylated form, the phosphoenolpyruvate: carbohydrate phosphoenolpyruvate transferase system (PTS) that can transport and concomitantly phosphorylate D-glucose were replaced with GalP, a D-galactose: H^+ symporter involved in D-glucose uptake (Luo et al., 2014). D-Glucose isomerase (DGI) (Liu et al., 2015; Jin et al., 2021), D-allulose 3-epimerase (DAE) (Kim et al., 2008), and ribose-5-phosphate isomerase (RPI) (Park et al., 2007; Yeom et al., 2011) were co-expressed to perform the reactions from D-glucose to D-allose, in which D-fructose and D-allulose were intermediates of the cascade epimerization. The kinases catalyzing the phosphorylation of D-glucose and D-fructose were knocked out, and the Embden-Meyerhof-Parnas (EMP) pathway was blocked, with the purpose to maximize the titer of the target product. In this work, our achievements provide an alternative strategy for D-allose synthesis, and may promote the research on fermentative production of rare sugars.

Materials and methods

Strains and media

The strains used in this study were listed in Table 1. *E. coli* JM109 (DE3) purchased from LMAI Bio (Shanghai) was the host for D-allose production. Luria–Bertani (LB) medium was composed of 10 g/L sodium chloride, 10 g/L tryptone, and 5 g/L yeast extract. The M9 medium for fermentation contained (per L): 4.78 g Na_2HPO_4, 2.99 g KH_2PO_4, 0.5 g NaCl, 0.12 g Mg_SO_4, 33 mg $CaCl_2$, 1 μg Thiamin, 1 μg Biotin, 50 mg EDTA, 0.1 mg H_3BO_3, 0.13 mg $CuCl_2$, 0.84 mg $ZnCl_2$, 8.3 mg $FeCl_3$, 0.1 mg $CoCl_2$, and 16 μg $MnCl_2$. D-Glucose and glycerol were added in M9 medium for D-allose synthesis and cell growth, respectively.

Plasmid construction

The plasmids of pRSFDuet-1 and pETDuet-1 for co-expressing protein were purchased from Novagen (Table 1). The genes of *dgi* (GenBank: AAA26838.1) (Jin et al., 2021), *dae* (GenBank: AAK88700.1) (Li et al., 2020), and *rpi* (GenBank:

TABLE 2 Primers applied in gene amplification and knockout.

Name	Primer sequence (5′-3′)	Gene
Gene amplification		
galP-F	CATGCCATGGCGATGCCTGACGCTAAAAAACAGGGG	*galP*
galP-R	CCAAGCTTCCTTAATCGTGAGCGCCTATTTCGCGC	
dgi-F	GGCATATGAACTACCAGCCGACCCC	*dgi*
dgi-R	CTCGAGTTAGCCACGGGCACCCAGC	
dae-F	CAATTGCATGAAGCACGGCATTTATTACAGC	dae
dae-R	GGGGTACCTTAGCCGCCCAGAACAAAACGG	
rpi-F	CGGGATCCGATGAAAATTGGCATTGGCAGCGATC	*rpi*
rpi-R	CGGAAGCTTTCACTTGCTGTACTTTTTTTCAATTTCGCC	
Gene deletion		
fruA-F	GGCATAATGAAAACGCTGCTGATTATTGACGCTAATCTCGGTCAGGCACGATTCCGGGGATCCGTCGACC	*fruA*
fruA-R	ATTACGCTGCTTTCGCTACTGCGTCCACTTCCGGACGTTTCAGGAAGGCAAGCGATTGTGTAGGCTGGAGCTGC	
ptsG-F	CCCATACTCAGGAGCACTCTCAATTATGTTTAAGAATGCATTTGCTAACCATTCCGGGGATCCGTCGACC	*ptsG*
ptsG-R	AGTCTCCCCAACGTCTTACGGATTAGTGGTTACGGATGTACTCATCCATCAGCGATTGTGTAGGCTGGAGCTGC	
pfkA-F	GTTCAGAGGTAGTCATGATTAAGAAAATCGGTGTGTTGACAAGCGGCGGTATTCCGGGGATCCGTCGACC	*pfkA*
pfkA-R	CGAAATCATTAATACAGTTTTTTCGCGCAGTCCAGCCAGTCACCTTTGAAAGCGATTGTGTAGGCTGGAGCTGC	
pfkB-F	CTGATTCGGTGCCAGACTGAAATCAGCCTATAGGAGGAAATGATGGTACGTATCATTCCGGGGATCCGTCGACC	*pfkB*
pfkB-R	GTTGGTGATGATTCCCCCAATGCTGGGGGAATGTTTTTGTTAGCGGGAAAGGAGCGATTGTGTAGGCTGGAGCTGC	
glk-F	CTTTAGCGGAGCAGTTGAAGAATGACAAAGTATGCATTAGTCGGTGATGTGGGCATTCCGGGGATCCGTCGACC	*glk*
glk-R	CCCGATATAAAAGGAAGGATTTACAGAATGTGACCTAAGGTCTGGCGTAAATGTGCAGCGATTGTGTAGGCTGGA GCTGC	
mak-F	CTACGCTATTGATATTGAAAAAAATAAGGAGAGTACCGTGCGTATAGGTATCATTCCGGGGATCCGTCGACC	*mak*
mak-R	CATGATGCGCCAATTGCCTACGTTTTTTACTCTTGTGGCCATAACCACGCAGCGATTGTGTAGGCTGGAGCTGC	
alsB-F	GCATCATCATCCGGCATCATTCAGTTTTATTGAGTGACCAGGATTGAATCAGCGATTGTGTAGGCTGGAGCTGC	*alsB*
alsB-R	CTCGGCAAGAATATTACAACTAACTTTGCTGGACGCTTTTTTTGTCTCTGATTCCGGGGATCCGTCGACC	

ABN53797.1) (Yeom et al., 2011) were optimized and synthesized by Beijing Genomics Institute (BGI, Beijing), which were then amplified by use of PrimeSTAR DNA polymerase (Takara). The genes of *dgi*, *dae*, *rpi*, and *galP* were cloned using primers *dgi*-F and *dgi*-R, *dae*-F and *dae*-R, *rpi*-F and *rpi*-R, *galP*-F and *galP*-R (Table 2), respectively, then digested with *Nde* I and *Xho* I, *Mfe* I and *kpn* I, *BamH* I and *Hind* III, *Nco* I, and *Hind* III (NEB, Beijing) at 37°C for 1 h. The obtained DNA fragments were ligated to pETDuet-1 or pRSFDuet-1 with the help of T4 DNA ligase (NEB, Beijing) at 16°C for 12 h, resulting in pETDuet-*dgi*, pRSFDuet-*dae*, pRSFDuet-*rpi*, pETDuet-*galP*-*dgi*, and pRSFDuet-*dae*-*rpi*, respectively.

Gene knockout

The genes of *fruA*, *ptsG*, *glk*, *mak*, *pfkA*, and *pfkB* located on the *E. coli* genome were knocked out by a λ red homologous recombination system (Datsenko and Wanner, 2000). The kan^R gene with two FRT sites were amplified by use of pKD13 as a template, and then electrotransferred to *E. coli* strains harboring pKD46. After the replacement of the target gene by the kan^R gene with the FRT sites, plasmid pCP20 was employed to express the DNA recombinase (FLP) for eliminating the kan^R gene (Baba et al., 2006; Dugar et al., 2016; Yang et al., 2018). The primers for gene knockout were shown in Table 2.

Enzyme analysis

E. coli strains were cultured in 100 ml LB medium with ampicillin (100 μg/ml) at 37°C and 220 rpm for 4 h. Protein expression was induced by the addition of 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) when the cell density (OD_{600}) reached ≈0.6. DGI, DAE, and RPI were expressed at 37°C and 220 rpm for 12 h, respectively. Then, *E. coli* strains were centrifuged at 8,000 *g* and 4°C for 10 min, which were washed with Tris-HCl (50 mM, pH 7.0) for two times, and finally suspended with 15 ml Tris-HCl (50 mM, pH 7.0). The cells were broken by ultrasonic crushing instrument JY92-IIN Jingxin co., ltd. (Shanghai) with ice-water bath for 5 min (3 s on and 3 s off). The cell debris was removed by centrifugation at 8,000 *g* and 4°C for 10 min. Protein in the supernatant was analyzed by use of 12% PAGE protein prefabricated gel

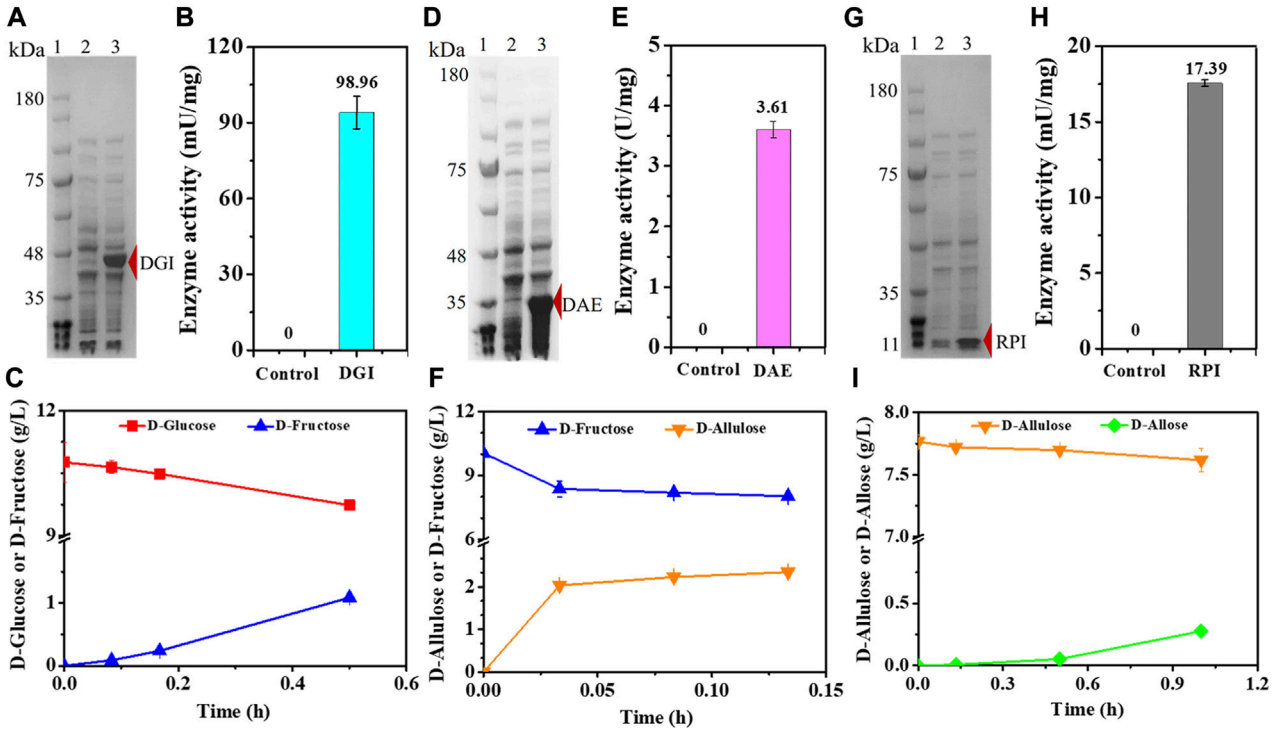


FIGURE 2
Expression and functionality of DGI, DAE, and RPI. **(A)** SDS-PAGE analysis of DGI. Lane 1: Marker, Lane 2: *E. coli* (Empty A), Lane 3: *E. coli* (DGI). **(B)** Activity of crude DGI. The supernatant of disrupted *E. coli* (Empty A) was used as a control. **(C)** Conversion of D-glucose to D-fructose by crude DGI. **(D)** SDS-PAGE analysis of DAE. Lane 1: Marker, Lane 2: *E. coli* (Empty B), Lane 3: *E. coli* (DAE). **(E)** Activity of crude DAE. The supernatant of disrupted *E. coli* (Empty B) was used as a control. **(F)** Conversion of D-fructose to D-allulose by crude DAE. **(G)** SDS-PAGE analysis of RPI. Lane 1: Marker, Lane 2: *E. coli* (Empty B), Lane 3: *E. coli* (RPI). **(H)** Activity of crude RPI. The supernatant of disrupted *E. coli* (Empty B) was used as a control. **(I)** Conversion of D-allulose to D-allose by crude RPI. Error bars indicated standard error (n = 3).

(KeyGEN Biotech, Nanjing). Proteins were analyzed by MicroSpectrophotometer K5500PLus from Kaiao Technology co. ltd. (Beijing). The activity analysis was carried out at 30°C in a reaction system (500 µl) containing 400 µl Tris-HCl buffer (50 mM, pH 7.0), proper sugar (10.8 g/L D-glucose, 10.1 g/L D-fructose, or 7.8 g/L D-allulose), and 100 µL crude enzyme. After the system was preheated at 30°C for 5 min, crude enzyme was added. The reaction was terminated by boiling at 2, 5, 8, 10, 15, 30, 60, 90, and 180 min, respectively. The supernatant was obtained by centrifugation at 6,000 *g* for 10 min, and then detected by high performance liquid chromatography (HPLC). One unit of DGI, DAE, or RPI activity (U/mg) was defined as the amount of producing 1 µmol of D-fructose, D-allulose or D-allose per minute, respectively.

Fermentation

E coli strains were incubated in 4 ml LB medium overnight at 37°C and 220 rpm, then cultured in 50 ml M9 medium containing appropriate D-glucose and glycerol with 100 µg/ml ampicillin and 50 µg/ml kanamycin. After the cell density (OD_{600}) reached 0.6, IPTG (0.2 mM) was added for induction, and the fermentation temperature was adjusted to 30°C. Fermentation samples were taken at an interval of 12 h. The cell density was measured with ultraviolet spectrophotometer. D-Glucose, glycerol, D-fructose, D-allulose, and D-allose were analyzed by a high-performance liquid chromatograph (HPLC, HITACHI) with a refractive index detector monitor. A column of Sugar-Pak™ I purchased from Waters was employed with a mobile phase (deionized water) flow rate of 0.5 ml/min at 85°C. The injection volume of the sample was 10 µl, with a retention time of 20 min. The peaks of D-glucose, glycerol, D-fructose, D-allulose and D-allose appeared at 9.7, 13.1, 11.3, 14.9, and 12.6 min, respectively. The target product D-allose was further verified by use of liquid chromatography-mass spectrometry (LC-MS) with Agilent 6,520 and Agilent 1,260 instruments. The conditions of LC were the same as HITACHI. The conditions of MS were as follows: mode, ESI (-); scan range, 100–310 m/z; capillary voltage, 3.5 kV; fragmentor voltage, 140 V; atomization pressure, 40 psi; gas, N_2; and gas temperature, 350°C. The standard substance of D-allose

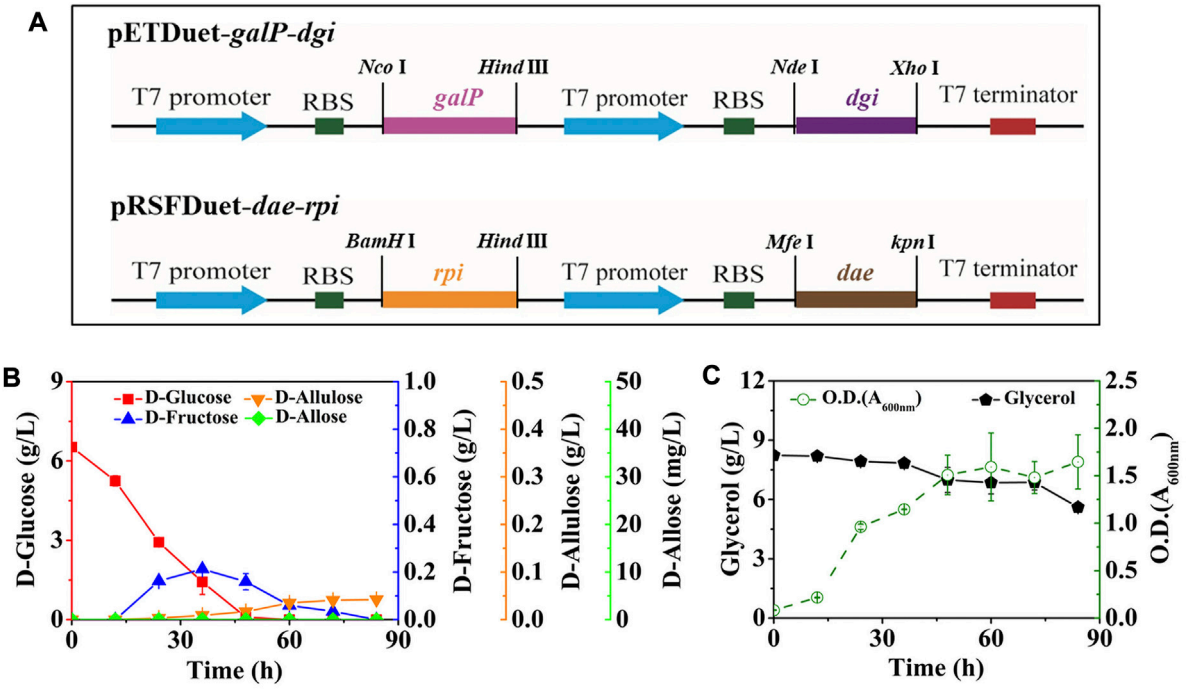


FIGURE 3
Synthesis of D-allose by fermentation using *E. coli* (DGI, DAE, RPI, GalP). *E. coli* (DGI, DAE, RPI, GalP) was cultured in M9 medium with D-glucose and glycerol at 30°C. **(A)** Recombinant plasmids of pETDuet-*galP-dgi* and pRSFDuet-*dae-rpi*. **(B)** D-Glucose, D-fructose, D-allulose, and D-allose. **(C)** Glycerol and cell density. Error bars indicated standard error (n = 3).

(purity ≥97%) was purchased from Yuanye Biotechnology (Shanghai).

Results and discussion

Rational design of an izumoring pathway for converting D-glucose to D-allose

As shown in Figure 1, the route designed for *in vivo* generation of D-allose was based on a three-step cascade of Izumoring reactions by use of D-glucose as a substrate. D-Glucose isomerase (DGI) was employed to convert D-glucose to D-fructose. This enzyme is also called D-xylose isomerase, and has been widely applied in industrial production of High-Fructose Corn Syrup (HFCS) (Liu et al., 2015), especially the DGI of *Streptomyces rubiginosus*. Although *S. rubiginosus* DGI exhibits good thermal stability, its activity is fully activated under alkaline conditions and rapidly decreases below pH 7.5, which is not conducive to its application in D-glucose isomerization in *E. coli* cells. Therefore, a *S. rubiginosus* DGI mutant with D56N and E221A was used in this work, whose catalytic efficiency has been reported to be significantly improved around neutral pH (Jin et al., 2021). The reaction from D-fructose to D-allulose can be carried out by D-tagatose 3-epimerase (DTE) or DAE. Compared with DTE, DAE has higher substrate specificity but lower thermal stability. Despite that, DAE is theoretically amenable to mild fermentation conditions, so we selected a DAE from *Agrobacterium tumefaciens*, which has been shown to be well expressed in *E. coli* (Kim et al., 2006). The synthesis of D-allose from D-allulose can be accomplished by L-rhamnose isomerase (LRhI), galactose-6-phosphate isomerase (GPI) or RPI, of which LRhI and GPI normally catalyze the formation of D-altrose (> 8%) as a byproduct (Yeom et al., 2011). Therefore, here we used a *Clostridium thermocellum* RPI with a R132E mutation, which has been reported to have a higher specific activity and catalytic efficiency for D-allulose than the wild-type enzyme (Lim and Oh, 2011; Yeom et al., 2011).

Expression and functionality of the enzymes involved in cascade epimerization

In order to confirm that the cascade epimerization from D-glucose to D-allose could be realized *in vivo*, we heterologously expressed DGI, DAE, and RPI in *E. coli*, respectively, resulting in

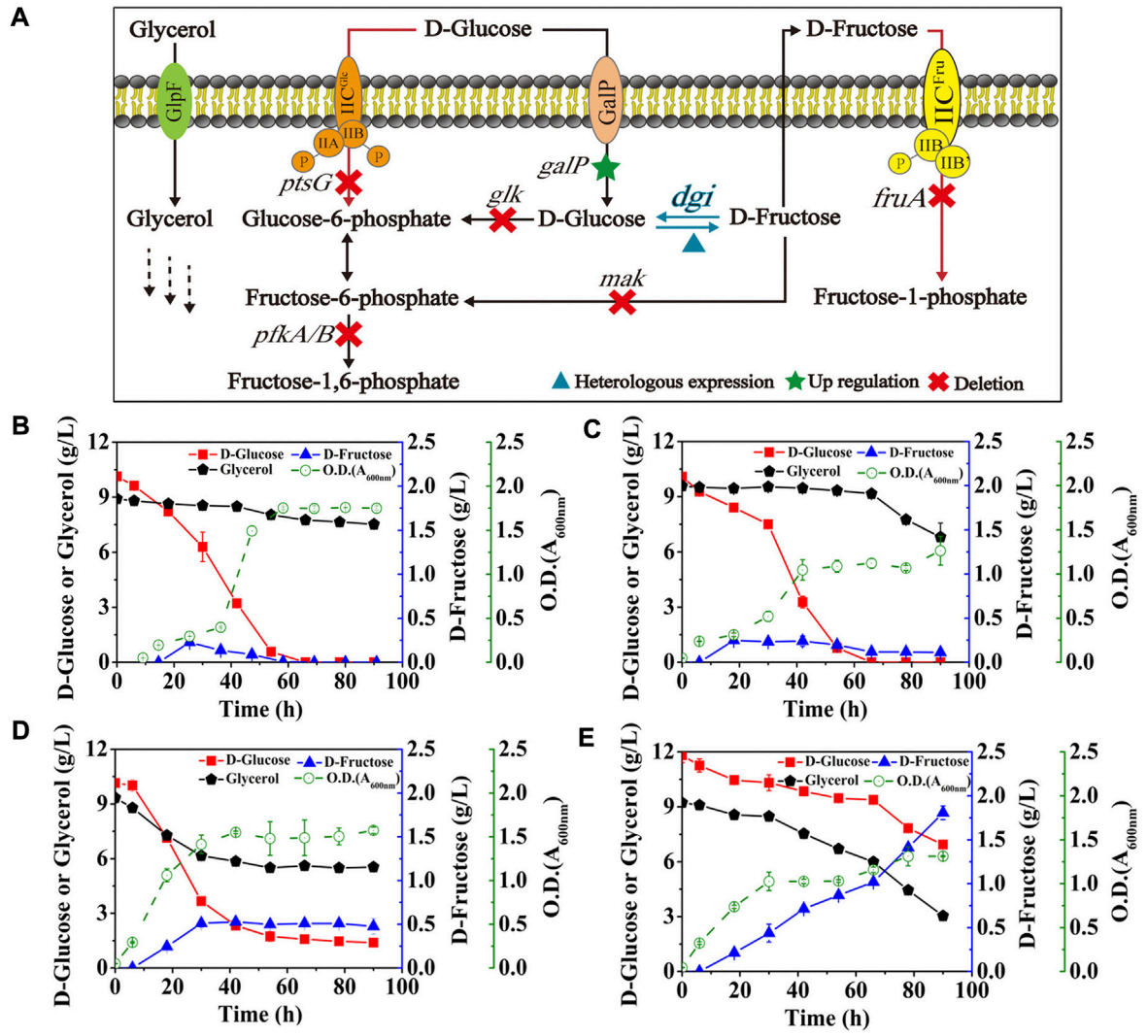


FIGURE 4
Effects of the phosphorylation acting on D-glucose and key intermediates. *E. coli* cells were cultured in M9 medium with D-glucose and glycerol at 30°C. **(A)** Schematic of the cell factory for D-fructose generation. **(B)** *E. coli* (DGI, GalP), **(C)** *E. coli* (DGI, GalP, ΔFruA), **(D)** *E. coli* (DGI, GalP, ΔFruA, ΔPtsG), **(E)** *E. coli* (DGI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, and ΔPfkB). Error bars indicated standard error (*n* = 3).

the strains *E. coli* (DGI), *E. coli* (DAE), and *E. coli* (RPI). Genes were expressed under T7 promoter with IPTG as an inducer at 37°C. Then, the catalytic function of each enzyme was tested. As shown in Figures 2A,D,G, the clear bands in SDS-PAGE indicate that DGI, DAE, and RPI could be produced in cells in soluble forms with molecular masses of around 45, 35, and 15 kDa, respectively, which are in good agreement with the values (43, 32, and 17 KDa) deduced from the amino acid sequences. When D-glucose (Figure 2B), D-fructose (Figure 2E) or D-allulose (Figure 2H) was used as a substrate, the crude DGI, DAE, and RPI were able to yield D-fructose, D-allulose, and D-allose in Tris-HCl buffer (50 mM, pH 7.4) at 30°C, with activities of 98.96 mU/mg, 3.61 U/mg, and 17.39 mU/mg, respectively. These data suggest that the designed route for cascade epimerization of D-glucose to D-allose by *E. coli* was basically feasible under fermentation conditions, and the conversion of D-allulose to D-allose should be the rate-limiting step since the expression level and the activity of RPI were both lower than those of DGI and DAE.

After that, we co-expressed DGI, DAE, RPI, and GalP, resulting in the strain *E. coli* (DGI, DAE, RPI, GalP). GalP is known as a D-galactose: H^+ symporter, which is also able to transport D-glucose in *E. coli*. This passage is not accompanied by phosphorylation (Luo et al., 2014), so that the D-glucose taken up by GalP can serve as a precursor for D-fructose synthesis. The mutant was cultivated in M9 medium containing 6.52 g/L

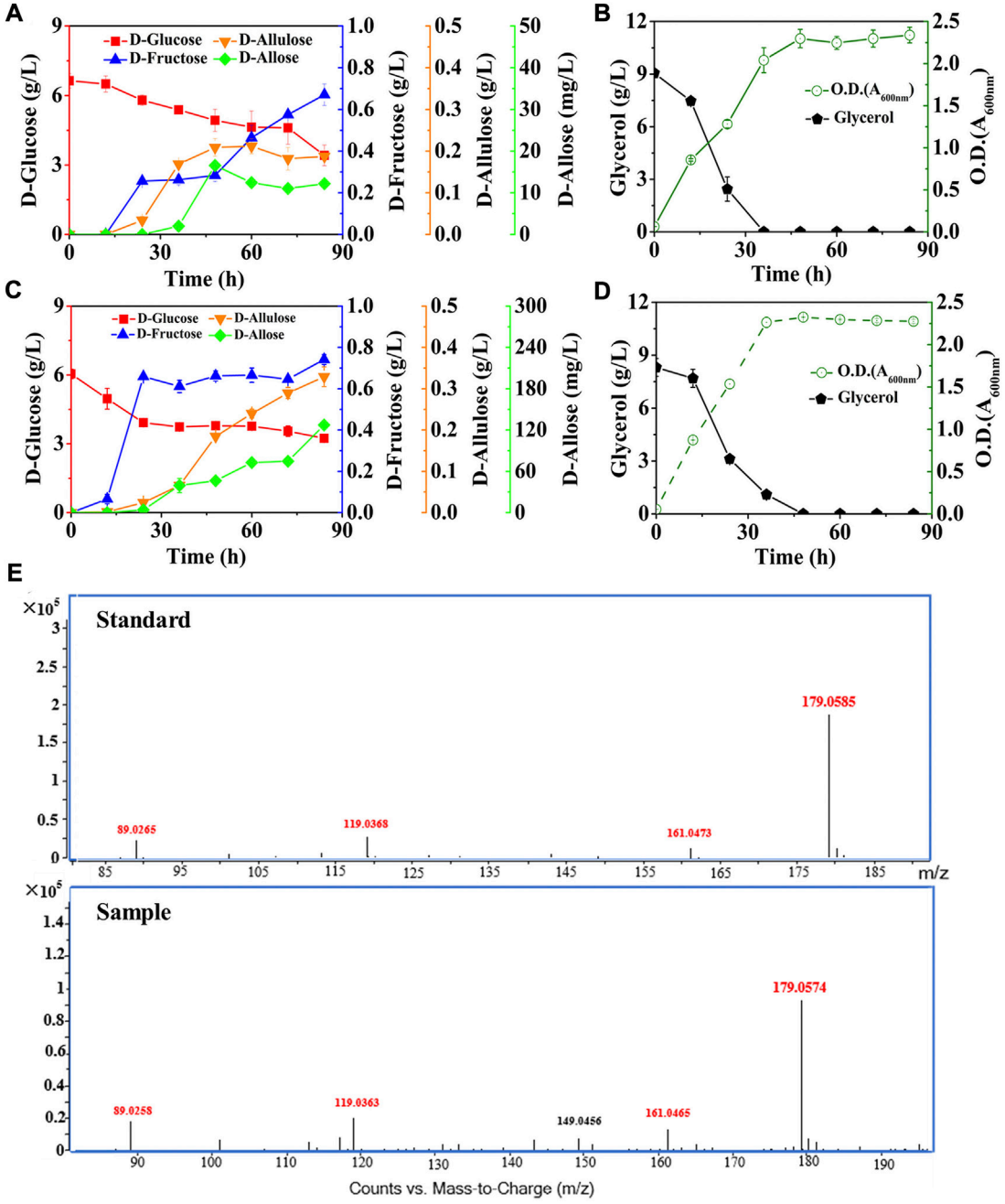


FIGURE 5
Deletion of D-allose transporter to improve the performance of cell factory. *E. coli* cells were cultured in M9 medium with D-glucose and glycerol at 30°C. **(A)** and **(B)** *E. coli* (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, and ΔPfkB). **(C)** and **(D)** *E. coli* (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, ΔPfkB, and ΔAlsB). **(E)** LC-MS analysis of D-allose standard and the fermentation sample by using *E. coli* (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, ΔPfkB, and ΔAlsB) at 60 h. Error bars indicated standard error (n = 3).

D-glucose and 8.23 g/L glycerol, with the aim of allowing cells to utilize D-glucose as a substrate for D-allose synthesis as much as possible, rather than a growth carbon source, thereby improving the yield of the product. The intermediates of D-fructose and D-allulose began to appear in the medium after 12 h of fermentation (Figure 3B). The D-fructose level first increased and then decreased, reaching a peak of 0.21 g/L at 36 h, while D-allulose gradually increased with time, reaching 0.04 g/L after 84 h. Unfortunately, we did not find the generation of the target product D-allose, and the utilization of glycerol might be limited

by carbon catabolite repression (CCR), since glycerol level remained unless D-glucose was depleted (Figure 3C).

Blocking of the phosphorylation acting on D-glucose and key intermediates

Because the activities of DGI, DAE, and RPI under fermentation conditions have been confirmed, we suspected that sugar phosphorylation might be the main reason for the failure of D-allose production, which has also been reported to be responsible for causing CCR (Fox and Prather, 2020). As an important intermediate, whether D-fructose could be continuously generated in cells was the key to synthesizing D-allose. We thus utilized the strain *E. coli* (DGI), which had a pathway capable of yielding D-fructose from D-glucose, to optimize the pathways related to the sugar phosphorylation in order to enhance D-fructose generation.

E. coli has been reported to take up D-glucose primarily through the glucose phosphoenolpyruvate transferase system (PTS) (Phue et al., 2005), which transports and concomitantly phosphorylates D-glucose to glucose-6-phosphate, thus allowing D-glucose to enter the EMP pathway as a carbon source for growth. However, expression of GalP in *E. coli* (DGI) did not obviously improve D-fructose production and glycerol co-utilization (Figure 4B) when compared with the data in Figure 3. Fructose PTS is the major route for uptake of D-fructose by *E. coli*. It is able to phosphorylate D-fructose to fructose-1-phosphate, followed by entering the EMP pathway as a carbon source, which can explain that the produced D-fructose would be gradually consumed in the middle and late stages of fermentation (Figure 4B). Also, CCR still occurs in theory due to the presence of fructose PTS, thus inhibiting the use of glycerol by cells. Therefore, we then deleted the gene of *fruA* involved in the fructose PTS, resulting in the strain *E. coli* (DGI, GalP, ΔFruA). As illustrated in Figure 4C, inactivation of fructose PTS effectively reduced the consumption of D-fructose, which could be maintained at 0.11 g/L even after 90 h. Meanwhile, the consumption of glycerol by *E. coli* (DGI, GalP, ΔFruA) was increased when compared with that of *E. coli* (DGI, GalP), which might be because the extracellular D-fructose could not be utilized without FruA, forcing the cells to use glycerol as a carbon source. We observed that the ability of *E. coli* (DGI, GalP, ΔFruA) to metabolize glycerol was severely inhibited by D-glucose, suggesting that the expression of GalP was not able to relieve glucose CCR. Therefore, the gene of *ptsG* involved in the glucose PTS was further deleted, resulting in the strain *E. coli* (DGI, GalP, ΔFruA, ΔPtsG). The data in Figure 4D show that D-glucose consumption for growth was reduced and glycerol utilization was further increased after the glucose PTS was inactivated. More importantly, D-fructose level was increased nearly 4-fold to approximately 0.48 g/L. However, the yield of D-fructose on D-glucose was still quite low, which was less than 0.06 g/g. We then further knocked out the genes of *glk*, *mak*, *pfkA*, and *pfkB* to block the phosphorylation of D-glucose and D-fructose in the cytoplasm and the EMP pathway (Figure 1). As a result, the obtained strain had a significant improvement in D-fructose synthesis, with a D-fructose level of 1.81 g/L and a yield of 0.37 g/g on D-glucose (Figure 4E).

Regulation of D-allose transport to enhance the forward epimerization reactions

Next, we co-expressed DAE and RPI in the optimized strain, resulting *E. coli* (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, and ΔPfkB), and cultivated the mutant in the M9 medium supplemented with around 6.00 g/L D-glucose and 9.00 g/L glycerol (Figures 5A,B). The data show that D-allose started to appear after 24 h of fermentation, and reached its maximal level of 16.57 mg/L at 48 h, then decreased to around 12.19 mg/L. Similarly, the intermediate D-allulose also had a slight decrease in the late stage of fermentation, finally reaching 0.19 g/L at 84 h. At the same time, the D-fructose produced was 0.67 g/L, while the consumption of D-glucose was about 3.21 g/L. These data indicate that the total level of the target and intermediate products was less than 0.99 g/L when fermentation was completed, which was much lower than the amount of the substrate consumed. There are no differences in the molecular masses of D-glucose, D-fructose, D-allulose and D-allose. Thus, it suggests that most of the consumed D-glucose flowed to other pathways rather than the designed pathway for yielding D-allose, probably due to the existence of unknown pathways in *E. coli* that could bypass the genes we deleted to utilize the hexoses involved in D-allose synthesis.

Here, the three reactions involved in the Izumoring cascade epimerization were all reversible, so one of the ways capable of further improving D-allose production was to make the effluxed D-allose unable to enter the mutant cells, thereby continuously driving the reversible reactions to the forward direction. It has been reported that the D-allose transport system of *E. coli* is composed of AlsA, AlsB, and AlsC (Kim et al., 1997; Poulsen et al., 1999), wherein AlsB is responsible for the specific binding of D-allose, followed by the transmembrane passage with the aid of AlsA and AlsC. We thus knocked out the gene of *alsB* to damage the D-allose transport system of *E. coli* (DGI, DAE, RPI, GalP, ΔFruA, ΔPtsG, ΔGlk, ΔMak, ΔPfkA, and ΔPfkB), and cultivated the mutant in M9 medium under similar conditions to our previous experiments (Figures 5C,D). It is observed that the D-allose produced, which was further confirmed by high-performance liquid chromatography-mass spectrometry (LC-MS) (Figure 5E), was 10-fold higher than that of the strain without AlsB deletion (Figure 5A), reaching 127.35 mg/L after 84 h, with a yield of over 0.045 g/g on D-glucose. More excitingly, the level of D-allose rose dramatically in the middle and late

stages of fermentation, and maintained an obvious upward trend even at the end of the experiment. To the best of our knowledge, this work is the first report of D-allose production from D-glucose by microbial fermentation. However, it is similar to enzymatic methods that the D-allose yield through fermentation was quite low. In this work, three reversible isomerization reactions were involved in the pathway synthesizing D-allose, which might seriously affect the yield of D-allose on D-glucose. In the future, we plan to introduce the phosphorylation-dephosphorylation pathway to convert D-glucose into D-allulose (Li et al., 2021), and then use RPI to isomerize D-allulose into D-allose, so as to further optimize the biosynthetic pathway.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

L-HF and H-DZ provided financial support and research program. L-JZ and QG were responsible for the specific experiments and data analysis. L-HF, L-JZ, and QG completed the first draft of the manuscript. Y-XZ and C-YL participated in data collection and manuscript revision. H-DZ reviewed the manuscript. All authors have played an important role in this work and approved it to be published.

Funding

This study was funded by the National Natural Science Foundation of China (No. 22278083 and 21978014), and Quanzhou City Science and Technology Program of China (No. 2021CT006).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* 2 (1), 2006.0008. doi:10.1038/msb4100050

Chen, Z., Chen, J., Zhang, W., Zhang, T., Guang, C., and Mu, W. (2018). Recent research on the physiological functions, applications, and biotechnological production of D-allose. *Appl. Microbiol. Biotechnol.* 102 (10), 4269–4278. doi:10.1007/s00253-018-8916-6

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97 (12), 6640–6645. doi:10.1073/pnas.120163297

Dugar, G., Svensson, S. L., Bischler, T., Waldchen, S., Reinhardt, R., Sauer, M., et al. (2016). The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in *Campylobacter jejuni*. *Nat. Commun.* 7, 11667. doi:10.1038/ncomms11667

Fox, K. J., and Prather, K. L. (2020). Carbon catabolite repression relaxation in *Escherichia coli*: Global and sugar-specific methods for glucose and secondary sugar co-utilization. *Curr. Opin. Chem. Eng.* 30, 9–16. doi:10.1016/j.coche.2020.05.005

Granström, T. B., Takata, G., Tokuda, M., and Izumori, K. (2004). Izumoring: A novel and complete strategy for bioproduction of rare sugars. *J. Biosci. Bioeng.* 97 (2), 89–94. doi:10.1016/s1389-1723(04)70173-5

Guggisberg, A. M., Park, J., Edwards, R. L., Kelly, M. L., Hodge, D. M., Tolia, N. H., et al. (2014). A sugar phosphatase regulates the methylerythritol phosphate (MEP) pathway in malaria parasites. *Nat. Commun.* 5, 4467–4477. doi:10.1038/ncomms5467

Guo, Q., Zheng, L. J., Luo, X., Gao, X. Q., Liu, C. Y., Deng, L., et al. (2021). Engineering *Escherichia coli* for D-allulose production from D-fructose by fermentation. *J. Agric. Food Chem.* 69 (45), 13578–13585. doi:10.1021/acs.jafc.1c05200

Iga, Y., Nakamichi, K., Shirai, Y., and Matsuo, T. (2010). Acute and sub-chronic toxicity of D-allose in rats. *Biosci. Biotechnol. Biochem.* 74 (7), 1476–1478. doi:10.1271/bbb.100121

Ishihara, Y., Katayama, K., Sakabe, M., Kitamura, M., Aizawa, M., Takara, M., et al. (2011). Antioxidant properties of rare sugar D-allose: Effects on mitochondrial reactive oxygen species production in Neuro2A cells. *J. Biosci. Bioeng.* 112 (6), 638–642. doi:10.1016/j.jbiosc.2011.08.005

Izumori, K. (2006). Izumoring: A strategy for bioproduction of all hexoses. *J. Biotechnol.* 124 (4), 717–722. doi:10.1016/j.jbiotec.2006.04.016

Jin, L. Q., Jin, Y. T., Zhang, J. W., Liu, Z. Q., and Zheng, Y. G. (2021). Enhanced catalytic efficiency and thermostability of glucose isomerase from *Thermoanaerobacter ethanolicus* via site-directed mutagenesis. *Enzyme Microb. Technol.* 152, 109931. doi:10.1016/j.enzmictec.2021.109931

Kim, C., Song, S., and Park, C. (1997). The D-allose operon of *Escherichia coli* K-12. *J. Bacteriol.* 179 (24), 7631–7637. doi:10.1128/jb.179.24.7631-7637.1997

Kim, H. J., Hyun, E. K., Kim, Y. S., Lee, Y. J., and Oh, D. K. (2006). Characterization of an *Agrobacterium tumefaciens* D-psicose 3-epimerase that converts D-fructose to D-psicose. *Appl. Environ. Microbiol.* 72 (2), 981–985. doi:10.1128/AEM.72.2.981-985.2006

Kim, N. H., Kim, H. J., Kang, D. I., Jeong, K. W., Lee, J. K., Kim, Y., et al. (2008). Conversion shift of D-fructose to D-psicose for enzyme-catalyzed epimerization by addition of borate. *Appl. Environ. Microbiol.* 74 (10), 3008–3013. doi:10.1128/AEM.00249-08

Li, C., Gao, L., Du, K., Lin, H., Ren, Y., Lin, J., et al. (2020). Production of D-allose from D-fructose using immobilized L-rhamnose isomerase and D-psicose 3-epimerase. *Bioprocess Biosyst. Eng.* 43 (4), 645–653. doi:10.1007/s00449-019-02262-y

Li, Y., Shi, T., Han, P., and You, C. (2021). Thermodynamics-driven production of value-added D-allulose from inexpensive starch by an *in vitro* enzymatic synthetic biosystem. *ACS Catal.* 11 (9), 5088–5099. doi:10.1021/acscatal.0c05718

Lim, Y. R., and Oh, D. K. (2011). Microbial metabolism and biotechnological production of D-allose. *Appl. Microbiol. Biotechnol.* 91 (2), 229–235. doi:10.1007/s00253-011-3370-8

Liu, Z. Q., Zheng, W., Huang, J. F., Jin, L. Q., Jia, D. X., Zhou, H. Y., et al. (2015). Improvement and characterization of a hyperthermophilic glucose isomerase from *Thermoanaerobacter ethanolicus* and its application in production of high fructose corn syrup. *J. Ind. Microbiol. Biotechnol.* 42 (8), 1091–1103. doi:10.1007/s10295-015-1639-0

Luo, Y., Zhang, T., and Wu, H. (2014). The transport and mediation mechanisms of the common sugars in *Escherichia coli*. *Biotechnol. Adv.* 32 (5), 905–919. doi:10.1016/j.biotechadv.2014.04.009

Miyawaki, Y., Ueki, M., Ueno, M., Asaga, T., Tokuda, M., and Shirakami, G. (2012). D-allose ameliorates cisplatin-induced nephrotoxicity in mice. *Tohoku J. Exp. Med.* 228 (3), 215–221. doi:10.1620/tjem.228.215

Mooradian, A. D., Smith, M., and Tokuda, M. (2017). The role of artificial and natural sweeteners in reducing the consumption of table sugar: A narrative review. *Clin. Nutr. ESPEN* 18, 1–8. doi:10.1016/j.clnesp.2017.01.004

Morimoto, K., Suzuki, T., Ikeda, H., Nozaki, C., and Goto, S. (2022). One-pot multi-step transformation of D-allose from D-fructose using a co-immobilized biocatalytic system. *J. Gen. Appl. Microbiol.* 68 (1), 2021.07.002–9. doi:10.2323/jgam.2021.07.002

Naha, N., Lee, H. Y., Mi, J. J., Chung, B. C., Kim, S. H., and Kim, M. O. (2008). Rare sugar D-allose induces programmed cell death in hormone refractory prostate cancer cells. *Apoptosis* 13 (9), 1121–1134. doi:10.1007/s10495-008-0232-7

Nakamura, T., Tanaka, S., Hirooka, K., Toyoshima, T., Miyamoto, O., Tamiya, T., et al. (2011). Anti-oxidative effects of D-allose, a rare sugar, on ischemia-reperfusion damage following focal cerebral ischemia in rat. *Neurosci. Lett.* 487 (1), 103–106. doi:10.1016/j.neulet.2010.10.004

Park, C. S., Yeom, S. J., Kim, H. J., Lee, S. H., Lee, J. K., Kim, S. W., et al. (2007). Characterization of ribose-5-phosphate isomerase of *Clostridium thermocellum* producing D-allose from D-psicose. *Biotechnol. Lett.* 29 (9), 1387–1391. doi:10.1007/s10529-007-9393-7

Phue, J. N., Noronha, S. B., Hattacharyya, R., Wolfe, A. J., and Shiloach, J. (2005). Glucose metabolism at high density growth of *E. coli* B and *E. coli* K: Differences in metabolic pathways are responsible for efficient glucose utilization in *E. coli* B as determined by microarrays and northern blot analyses. *Biotechnol. Bioeng.* 90 (7), 805–820. doi:10.1002/bit.20478

Poulsen, T. S., Chang, Y. Y., and Jensen, B. H. (1999). D-Allose catabolism of *Escherichia coli*: Involvement of *alsI* and regulation of *als* regulon expression by allose and ribose. *J. Bacteriol.* 181 (22), 7126–7130. doi:10.1128/JB.181.22.7126-7130.1999

Su, L., Sun, F., Liu, Z., Zhang, K., and Wu, J. (2018). Highly efficient production of *Clostridium cellulolyticum* H10 D-psicose 3-epimerase in *Bacillus subtilis* and use of these cells to produce D-psicose. *Microb. Cell Fact.* 17 (1), 188. doi:10.1186/s12934-018-1037-1

Sun, Y., Hayakawa, S., Puangmanee, S., and Izumori, K. (2006). Chemical properties and antioxidative activity of glycated α-lactalbumin with a rare sugar, D-allose, by Maillard reaction. *Food Chem. x.* 95 (3), 509–517. doi:10.1016/j.foodchem.2005.01.033

Vigo, E., Stortz, C. A., and Marino, C. (2022). D-Allose, a rare sugar. Synthesis of D-allopyranosyl acceptors from glucose, and their regioselectivity in glycosidation reactions. *Org. Biomol. Chem.* 20 (22), 4589–4598. doi:10.1039/D2OB00590E

Wang, Y., Carder, H. M., and Wendlandt, A. E. (2020). Synthesis of rare sugar isomers through site-selective epimerization. *Nature* 578 (7795), 403–408. doi:10.1038/s41586-020-1937-1

Yamaguchi, F., Kamitori, K., Sanada, K., Horii, M., Dong, Y., Li, S., et al. (2008). Rare sugar D-allose enhances anti-tumor effect of 5-fluorouracil on the human hepatocellular carcinoma cell line HuH-7. *J. Biosci. Bioeng.* 106 (3), 248–252. doi:10.1263/jbb.106.248

Yang, H., Lu, X., Hu, J., Chen, Y., Shen, W., and Liu, L. (2018). Boosting secretion of extracellular protein by *Escherichia coli* via cell wall perturbation. *Appl. Environ. Microbiol.* 84 (20), e01382-18. doi:10.1128/AEM.01382-18

Yeom, S. J., Seo, E. S., Kim, Y. S., and Oh, K. D. K. (2011). Increased D-allose production by the R132E mutant of ribose-5-phosphate isomerase from *Clostridium thermocellum*. *Appl. Microbiol. Biotechnol.* 89 (6), 1859–1866. doi:10.1007/s00253-010-3026-0

Yokohira, M., Hosokawa, K., Yamakawa, K., Saoo, K., Matsuda, Y., Yu, Z., et al. (2008). Potential inhibitory effects of D-allose, a rare sugar, on liver preneoplastic lesion development in F344 rat medium-term bioassay. *J. Biosci. Bioeng.* 105 (5), 545–553. doi:10.1263/jbb.105.545

Zhu, Y., Yan, M., Wei, B., Li, X. B., Ma, Y., Sun, Y., et al. (2012). Overexpression of D-psicose 3-epimerase from *Ruminococcus sp.* in *Escherichia coli* and its potential application in D-psicose production. *Biotechnol. Lett.* 34 (10), 1901–1906. doi:10.1007/s10529-012-0986-4

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 06 December 2022
DOI 10.3389/fbioe.2022.1051233

OPEN ACCESS

EDITED BY
Yinjie Tang,
Washington University in St. Louis,
United States

REVIEWED BY
Yanran Li,
University of California, Riverside,
United States
Daochen Zhu,
Jiangsu University, China
Wei-Qin Zhuang,
The University of Auckland,
New Zealand

*CORRESPONDENCE
Mingzhu Ding,
mzding@tju.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 22 September 2022
ACCEPTED 25 November 2022
PUBLISHED 06 December 2022

CITATION
Cao Z, Yan W, Ding M and Yuan Y (2022), Construction of microbial consortia for microbial degradation of complex compounds.
Front. Bioeng. Biotechnol. 10:1051233.
doi: 10.3389/fbioe.2022.1051233

Construction of microbial consortia for microbial degradation of complex compounds

Zhibei Cao[1,2], Wenlong Yan[1,2], Mingzhu Ding[1,2]* and Yingjin Yuan[1,2]

[1]Frontier Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [2]Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin University, Tianjin, China

Increasingly complex synthetic environmental pollutants are prompting further research into bioremediation, which is one of the most economical and safest means of environmental restoration. From the current research, using microbial consortia to degrade complex compounds is more advantageous compared to using isolated bacteria, as the former is more adaptable and stable within the growth environment and can provide a suitable catalytic environment for each enzyme required by the biodegradation pathway. With the development of synthetic biology and gene-editing tools, artificial microbial consortia systems can be designed to be more efficient, stable, and robust, and they can be used to produce high-value-added products with their strong degradation ability. Furthermore, microbial consortia systems are shown to be promising in the degradation of complex compounds. In this review, the strategies for constructing stable and robust microbial consortia are discussed. The current advances in the degradation of complex compounds by microbial consortia are also classified and detailed, including plastics, petroleum, antibiotics, azo dyes, and some pollutants present in sewage. Thus, this paper aims to support some helps to those who focus on the degradation of complex compounds by microbial consortia.

KEYWORDS

microbial consortia, degradation, complex compounds, plastic biodegradation, petroleum biodegradation

Introduction

In recent years, increasingly serious pollution has been a major threat to public health, and more and more people are putting forward higher requirements for environmental restoration. Microbial environmental remediation is cleaner and more economical than the traditional burning landfill, which causes subsequent environmental pollution. The use of microbial consortia to degrade various pollutants into non-toxic or less-toxic compounds is a better option (Azubuike et al., 2016). At present, research on

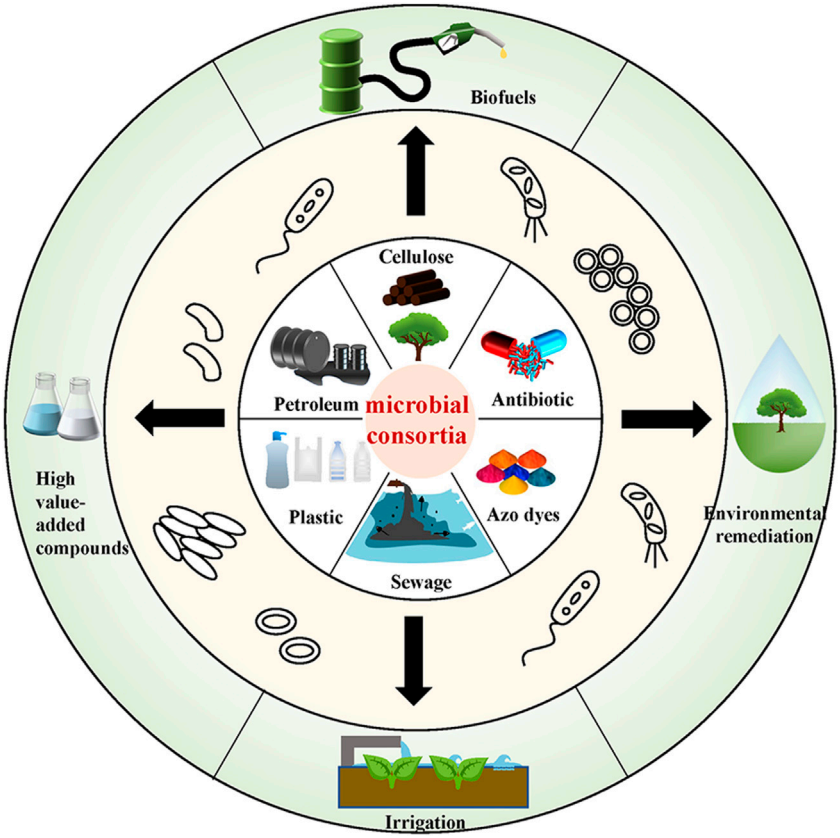


FIGURE 1
Microbial consortia bioremediation and reuse of complex compounds.

bioremediation by microbial consortia has practical significance, and it has already been applied in some cases. As shown in Figure 1, microbial consortia can degrade complex compounds, including plastics, petroleum, antibiotics, azo dyes, and some pollutants present in sewage. Furthermore, it also can be used in consolidated bioprocessing (CBP), which is a great solution to energy shortages. Petroleum hydrocarbons and plastics can also be used as raw materials for the production of high-value-added products. Because of the excellent degradation ability of microbial consortia for complex compounds, they are more commonly used than the single stain in environmental remediation. For example, they can well degrade the complex compounds in soil and sewage. Furthermore, shrubs and trees can be planted on the treated soil, and the treated sewage can be used as irrigation water for non-edible commercial crops (Biswas et al., 2021).

Creating an artificial microbial consortia system distributes the desired multiple catalytic enzyme expression pathways to different strains, and then co-culturing all strains to complete the task (Zhang and Stephanopoulos, 2016; Li Z. et al., 2019). The microbial consortia can degrade complex compounds that cannot be decomposed by a single bacterial system, such as starch and cellulose. Some complex compounds are difficult to be degraded due to their complex structures. However, some strains can break down these complex substrates into small-molecule sugars that can be used as carbon sources for other strains in the system (Wang S. et al., 2019; Tondro et al., 2020). For strains in the consortia, the rational division of metabolic pathways can reduce cross-reactions and thus the metabolic burden of each cell (Said and Or, 2017; Shen et al., 2020). Compared with a natural microbial consortia system, the composition of an artificial microbial consortia system is simpler, the division of labor is clearer, and it can be further modified for different target products (Qian et al., 2020; Zhang and Hong, 2020). Thus, a microbial consortium can be constructed to degrade a wide range of complex compounds precisely, which can enable the modular assembly and optimization of metabolic pathways by modulating the microbial consortia structure (Jones et al., 2017; Jones and Wang, 2018; Roell et al., 2019). Cross-feeding between bacteria can also be used to eliminate feedback inhibition and remove products or by-products, which is important for improving the degradation efficiency of complex compounds (Zhou et al.,

2015). Microbial consortia also have strong adaptability to, and stability within, complex environments (Kaeberlein et al., 2002). After a variety of cells with different functions are fused, the dynamic balance is maintained through complex interactions between cells, making the entire system more adaptable and stable when facing environmental fluctuations (McCarty and Ledesma-Amaro, 2019). The synergistic development of systems and synthetic biology will provide both a thorough understanding and a rational engineering of these complicated consortia for novel applications (Song et al., 2014).

The advantages described above are inspiring more and more researchers to explore the ability of microbial consortia to degrade complex compounds. In this paper, to provide a reference for the construction of microbial consortia, those that are currently used to degrade complex compounds are summarized, and future research directions for their construction are discussed.

Construction strategy of microbial consortia for degradation

Artificial microbial consortia systems have been developed and studied based on natural microbial consortia systems. There are usually two principles for the design of artificial microbial consortia systems: the top-down approach and the bottom-up approach. The top-down approach uses carefully selected environmental variables that force an existing microbiome (naturally occurring or inoculated) through ecological selection to perform the desired biological processes. This requires us to conceptualize the microbial consortia as a system model and determine the inputs and outputs of the system, including physical and chemical conditions, known abiotic and biological processes, environmental variables, and how operations on the microbial consortia promote or inhibit the biological processes being optimized (Lawson et al., 2019). The most commonly used method is to artificially enrich and screen functional microbial consortia. Although the conventional top-down approach offers a framework and has been widely successful for wastewater treatment and bioremediation, it often ignores processes that depend on intricate interactions between consortia members. Recent advances in synthetic biology have enabled researchers to develop bottom-up approaches and focus on engineering the microbiome's metabolic network and microbial interactions. The general design process is to obtain the genomes of individual members of the microbiome and then reconstruct the metabolic networks. The individual populations' reactions and metabolites can be compartmentalized and metabolic fluxes within and between populations can be simulated using optimality principles (Orth et al., 2010). These models can also simulate steady-state flux distributions over time and space. Such bottom-up tools provide a platform for rationally designing microbiomes with specific properties such as distributed pathways, modular species interactions, community resistance and resilience, and spatiotemporal organization that optimize ecosystem function and stability. Therefore, extending these designs to systems with non-model organisms of tens to hundreds of different species will require deeper insights into their metabolism and the principles governing their interactions and higher-order behavior (Lawson et al., 2019). Most microbial consortia that degrade complex compounds are constructed with the top-down approach.

When constructing microbial consortia to degrade complex compounds, one of the important issues is to select suitable chassis strains with suitable catalytic performance; whether they can coexist with other strains also needs to be considered (Jawed et al., 2019). Therefore, in the selection of chassis strains, strains with low mutation rates, non-toxic by-products, and high tolerance are generally selected. The next issue that needs to be considered is the division of degradation pathways. Long degradation pathways can be rationally divided into several strains, and different degradation pathways can be responsible for different strains (Lu et al., 2019). Although an artificial microbial consortia system can reduce the metabolic burden of cells, excessive segmentation of metabolic pathways will also lead to confusion and reduce the efficiency of mass transfer (Goers et al., 2014). In recent years, it has been discovered that the ordered spatiotemporal distribution of strains can improve the efficiency of microbial consortia to degrade complex compounds. In this way, each strain in the microbial consortia is provided with a suitable environment for degradation and a spatial position corresponding to the time sequence in the degradation pathway. Strain immobilization is a commonly used spatio-temporal distribution application, and plays an important role in promoting the biodegradation of complex compounds. It can be implemented through an ambient medium. Some researchers developed a special hydrogel as a new carrier to be used in the immobilization of artificial microbial consortia systems. This kind of hydrogel not only does not affect the material exchange of bacteria but also has a preservation effect on bacteria, which is conducive to the stability of their function. Strains with different environmental requirements in the microbial consortium can be preserved in different hydrogels, and the mixing of hydrogels does not change their individual properties (Johnston et al., 2020). The design of the culture device is also helpful to the spatio-temporal distribution and control of strains in an artificial microbial consortia system. Microfluidic technology achieves the fine regulation of different strains and improves the control of the microbial consortia system (Wang C. et al., 2019). Some researchers designed a ventilated biofilm reactor based on the gradient distribution of oxygen in space to achieve the reasonable coexistence and functional complementarity of three kinds of bacteria, which

effectively improved the efficiency of the microbial consortia (Shahab et al., 2020).

During the degradation of complex compounds, the carbon source required for the growth of microbial consortia is generally a complex compound itself, but not all microorganisms in the microbial consortia can utilize complex compounds as carbon sources. A common solution is to construct sequential utilization patterns of substrates and intermediates. Applying this model can not only avoid substrate competition but also eliminate the negative feedback inhibition caused by some by-products (Park et al., 2020). However, this sequential utilization pattern does not have well-defined material and energy flow paths like most microbial consortia for synthesize compounds *de novo*. The material and energy flow pathways in microbial consortia for degrading complex compounds are more reticular in structure. Material and energy are transferred repeatedly between strains and may in any case be consumed rather than eventually pooling in a product. This complex interaction network of material and energy is beneficial for degrading complex compounds because it can make the structure of the microbial consortium more stable and more resistant to environmental fluctuations. In this microbial consortium, the relationship between strains becomes more complex as the number of strains increases. For microbial consortiums with many strains, it might be critical to consider higher-order interactions (HOIs) to ensure stable coexistence and function (Mayfield and Stouffer, 2017). For example, in a three-member consortium, a third population could attenuate the negative interaction between two antagonistic populations. The consortia can also be stable even if the third species is antagonistic to the two species so long as each population modulates the inhibitory interactions between the remaining two members (Kelsic et al., 2015). Thus, the presence of an additional population could synergize with an existing community, resulting in a more stable consortium. The HOIs can also extend to more population network topologies (Grilli et al., 2017). Furthermore, the importance of HOIs increases with the number of populations in the microbial consortium (Friedman et al., 2017).

Cross-feeding and quorum sensing (QS) are two commonly used artificial design approaches to maintain complex stability. Symbiotic relationships in microbial consortia with few strains are primarily based on single metabolite cross-feeding, such as an amino acid (Harcombe et al., 2018). For example, amino acid auxotrophies can create complex interdependencies between microorganisms. These relationships promote stability and robustness by allowing for metabolic redundancy among community members (Embree et al., 2015). However, the secretion of a single metabolite is often insufficient to support the normal growth of all strains in a big microbial consortium, limiting its robustness and stability. In a microbial consortium degrading complex compounds with many strains, developing a multiple-metabolite cross-feeding strategy is closer to the reality, which is used to strengthen the correlation between microbial entities. Central to this strategy is the selection of appropriate metabolic branches for cross-feeding, which involve multiple metabolites that are critical for cell growth and translocate across cell membranes. Amino acid anabolism and energy metabolism can often be selected to establish close cell–cell correlations resulting in a very stable co-culture system (Li et al., 2022). The social and gregarious behavior of single-celled organisms such as bacteria is usually accomplished through intercellular communication, which can occur through QS. QS primarily regulates collective features that involve energetically costly "public goods" and are most effective or even only functional if performed by a microbial consortium. Bacterial traits controlled by QS include genetic phenotypes, biofilm formation, promoting or inhibiting function, and virulence (Mashruwala et al., 2022; Pütz et al., 2022; Ramsay et al., 2022). QS even can be a driver and target of other functions (Striednig and Hilbi, 2022). However, one challenge in incorporating more members within microbial consortia that degrade complex compounds is that many QS systems are not completely orthogonal, and one solution is to design a new QS system. Recently, a sophisticated QS circuit with high dynamic ranges, low leakiness, and the ability to simultaneously regulate multiple sets of genes in 1 cell was designed and was used to autonomously and temporally regulate three metabolic fluxes involved in a pathway (Ge et al., 2022). This was a big step forward but not sufficient to deal with the more complex situation in microbial consortia. It was also discovered that QS systems can be used for cell–cell communication between distant populations (Luo X. et al., 2015). QS systems may play a key role in microbial consortia that degrade complex compounds, just as they now play an important role in synthetic microbial consortia with fewer strains, but they must be studied further.

Current status of biodegradation of complex compounds by microbial consortia

Complex compounds are usually difficult to be efficiently degraded by natural microorganisms due to the complexity of their structures. Many researchers use microbial consortia to degrade complex compounds, especially common environmental pollutants. The current research progress on the degradation of complex compounds by microbial consortia is shown in Table 1. Common types of waste plastics such as polyethylene terephthalate (PET), polyethylene (PE), polystyrene (PS), and polyurethane (PU) have been degraded by microbial consortia. The study of strain interaction in natural microbial consortia is a necessary prerequisite for their construction. In fact, some isolated natural microbial consortia have the ability to degrade plastics. On this basis, researchers can add microorganisms to an isolated microbial consortium according to the relationships of the consortium to improve its efficiency, or build a simple microbial consortium to better understand degradability, gene

TABLE 1 An overview of recent advances in the degradation of complex compounds by microbial consortia.

Substrate		Achievement	Co-culture strains	References
Plastic	PET	the weight loss of PET film reached 23.2% in 7 days	*Rhodococcus*, *Pseudomonas putida*, and two metabolically engineered *Bacillus subtilis* species	Qi et al. (2022)
	PE	demonstrated 81% ± 4% of weight reduction for LDPE strips over a period of 120 days	*Enterobacter* sp. bengaluru-btdsce01, *Enterobacter* sp. bengaluru-btdsce02, and *Pantoea* sp. bengaluru-btdsce03	Bardají et al. (2020)
	PS	demonstrated 12.4% of weight reduction PS, a weight loss of 23% of HIPS film in 30 days	*Bacillus spp*. and *Pseudomonas spp*.	Ho et al. (2018)
	PU	50.3% of proprietary aromatic PE-PU-A copolymer was consumed in 25 days	*Rhodobacterales, Rhizobiales, Burkholderiales, Actinomycetales, Sphingobacteriales*	Gaytán et al. (2020)
Petroleum hydrocarbons	n-alkane	synergistic rate of biodegradation of diesel oil was 85.54% ± 6.42%	*Pseudomonas stutzeri, Dietzia sp.*	Hu et al. (2020)
	Polycyclic aromatic hydrocarbons	nearly completely degraded fluorene and phenanthrene after 5 days	*Sphingomonas, Pseudomonas, Sphingobium, Dokdonella* and *Luteimonas*	Bacosa and Inoue, (2015)
Antibiotic		78.3% of Sulfonamide antibiotics had degraded after 4 weeks	*Firmicutes* and *Bacteroides*, represented by *Bacillus* and *Flavobacterium*	Liao et al. (2016b)
		All sulfamethoxazole (5 mg/L) had degraded in 3 h	nitrifying sludge (e.g., *Nitrosomonas, Dokdonella, Defluviicoccus, Pseudomonas, Zoogloea, Thauera,* and *Pseudomonas*)	Yan et al. (2022)
		After 28-day incubation at 25°C, the ciprofloxacin loss was nearly 100%	Classes *Gammaproteobacteria, Bacteroidia, Betaproteobacteria* and *Leucobacter*	Liao et al. (2016a)
Azo dyes		98.2% decolorization	A halophilic bacterial consortium from textile wastewater	Shi et al. (2021)
		All Orange II (250 mg/L) had been decolorization	*Vanrija humicola, Meyerozyma caribbica, Debaryomyces hansenii,* and *Meyerozyma guilliermondii*	Samir Ali et al. (2022)
Wastewater		NH_4-N removal (100%) was observed within 7 days	*Chlorella vulgaris* and nitrifier-enriched-activated-sludge	Sepehri et al. (2020)
		the removal efficiency of acetoacetanilide (3200 mg/L) achieved 69.28 ± 0.42% within 14 days	*Paenarthrobacter, Rhizobium, Rhodococcus, Delftia* and *Nitratireductor*	Zhang et al. (2023)
Lignocellulose		produced 3.94 g/L butanol, which was five times higher than the control	C. *cellulovorans* and C. *beijerinckii*	Wen et al. (2020)
		7.61 g/L of butanol was generated from untreated corncob	*Thermoanaerobacterium thermosaccharolyticum* and *Clostridium acetobutylicum*	Jiang et al. (2020)
		0.44 g/g bioethanol production in biological pretreatment of the lignocellulosic cotton stalk	*Saccharomyces cerevisiae* YPH499 and *Pachysolen tannophilus* 32691	Malik et al. (2021)

regulation, or enzymatic activities. For example, about 90% of n-alkanes and aromatic hydrocarbons in petroleum hydrocarbons can be degraded by reconstructed microbial consortia (Bacosa and Inoue, 2015; Hu et al., 2020). Another obvious improvement is the treatment of sewage microbial consortia. Adding microalgae to activated sludge (a natural microbial consortium) can improve the adsorption and degradation efficiency of various compounds and even heavy metals in sewage (Sepehri et al., 2020). Alternatively, a microbial consortium can be constructed based on the interaction relationship between microorganisms, whose substrates are more targeted and generally can only degrade one or several specific types of complex compounds. At the same time, fewer strains are needed in the consortium, as each has its own clear mission. In artificial microbial consortia, the degradation process can be clearly represented and easily studied and even regulated. A representative example is the artificial two-strain consortium for CBP. One strain degrades complex cellulose into small molecular compounds, and another strain uses those small molecular compounds as substrates to synthesize the desired products (Wen et al., 2020). For complex compounds, especially pollutants in the environment (due to their refractory degradation and the complexity of the environment), the microbial consortium is one of the best choices for bioremediation.

Degradation of petroleum hydrocarbons by microbial consortia

Petroleum hydrocarbons are important energy resources and raw materials for all walks of life. Petroleum hydrocarbon

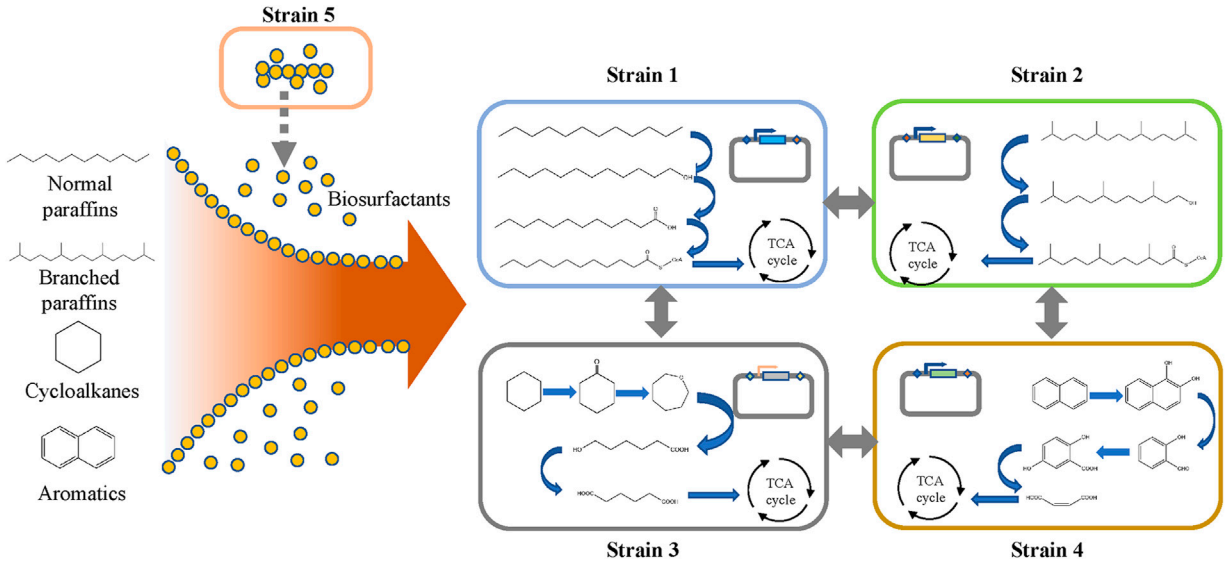


FIGURE 2
Biodegradation of petroleum hydrocarbons.

pollutants, such as normal paraffin, cycloalkanes, and aromatics, are recalcitrant compounds and are listed as priority pollutants (Holliger et al., 1997; Costa et al., 2012; Sajna et al., 2015; Li et al., 2020; Qian et al., 2021). They can usually be degraded in the presence of several natural microorganisms, each of which can decompose a specific set of molecules. Thus, microbial consortia have advantages in crude oil bioremediation (Zanaroli et al., 2010). Microbes in oil-contaminated areas adapt to the environment, resulting in genetic mutations in offspring that enable them to degrade petroleum hydrocarbon compounds (McDonald et al., 2006; Varjani and Upasani, 2016). Many novel species of microorganisms such as *Anaerobaculum*, *Desulfacinum infernum*, *Methanococcus thermolithotrophicus*, *Thauera phenylacetica*, and *Geobacillus subterraneus* have been isolated from sites of contamination (Garcia and Oliveira, 2013).

The top-down approach is often used to construct microbial consortia for the degradation of petroleum hydrocarbons. The degradation efficiency can be improved or the substrate range can be broadened by adding new strains to natural microbial consortia. The bioremediation capacity of microbial consortia in oil-contaminated areas is often limited due to the poor biodiversity of native microbial consortia, where the presence of microorganisms with complementary substrate specificities to degrade different hydrocarbons is lacking (Ron and Rosenberg, 2014). Microbial consortia with the potential to degrade petroleum hydrocarbon compounds can be screened out from oil-contaminated areas, before improving the degradation ability or substrate extensiveness of the microbial consortia system by artificial compounding or adding artificially engineered bacteria. There are reports showing that microbial consortia are superior to single bacteria in utilizing hydrocarbon contaminants in petroleum crude oil as the sole carbon source (Varjani et al., 2013). Such consortia show an increased degradation rate of diesel and polycyclic aromatic hydrocarbons (PAHs) when cultured under laboratory conditions (Varjani and Upasani, 2013). Ibrar and Zhang (2020) constructed a microbial consortium containing *Lysinibacillus, Paenibacillus, Gordonia,* and *Cupriavidus spp*. that could produce biosurfactants to enhance the ability of other bacteria to degrade petroleum hydrocarbons. Their results showed that the microbial consortia could use common polycyclic aromatic hydrocarbon pollutants (naphthalene and anthracene) as the sole carbon source. Therefore, artificial microbial consortia systems are a potential research direction to improve bioremediation efficiency in oil-contaminated areas (Varjani et al., 2015).

The bottom-up approach can also be used to construct microbial consortia based on the degradation pathway of petroleum hydrocarbons, which has been elucidated in the literature. As shown in Figure 2, the biodegradation of petroleum hydrocarbon can be divided into several processes (Li and Ding, 2021). In the first step, microorganisms enhance the bioavailability of petroleum hydrocarbon pollutants by chemotactic movements and secreting surfactants (Ahmad et al., 2020). These surface-active materials increase the surface area and bioavailability of hydrophobic and water-insoluble substrates, thereby increasing the speed at which petroleum hydrocarbons can approach microorganisms. Then, the petroleum hydrocarbons enter the cell through the transport process, mainly by free diffusion, passive transport, active

transport, and endocytosis (Gu et al., 2016). Finally, the petroleum hydrocarbon is degraded in the cell. The degradation pathways of petroleum hydrocarbon compounds mainly include aerobic degradation and anaerobic degradation. Common pathways for the degradation of linear alkanes include the initial degradation of alkanes and the oxidation of methyl groups, leading to the formation of alcohols, followed by the dehydrogenation of aldehydes to form their corresponding carboxylic acids. Then, the fatty acids are metabolized by the β-oxidation pathway (Abbasian et al., 2015). The degradation of cycloalkanes and aromatic hydrocarbons is more difficult than that of linear alkanes, as the former needs to be sequentially opened by hydrolase or isomerase and then degraded through degradation pathways that are different from those of linear alkanes (Gupta et al., 2015; Ghosal et al., 2016; Dhar et al., 2020; Li and Ding, 2021). Furthermore, each degradation pathway of petroleum hydrocarbons is relatively long, which will bring greater growth pressure to cells. Under the condition of ensuring degradation efficiency, a single strain cannot undertake all the functions of petroleum hydrocarbon degradation and biosurfactant production at the same time. Therefore, many studies are using microbial consortia to degrade petroleum hydrocarbons, where different bacterial species undertake different functions in the degradation process. This can not only reduce the growth pressure of individual cells but also improve the tolerance of bacterial groups to harsh environments through cooperation between different bacterial species, thereby making the entire degradation system stable and robust.

The degradation efficiency of petroleum hydrocarbons can be improved by modifying engineered bacteria according to the degradation pathway. Such genetically modified engineered microorganisms can degrade or assist in the degradation of complex compounds. Luo Q. et al. (2015) constructed oil biodegradation bacteria to promote the biodegradation of diesel. The alkane hydroxylase (alkB) gene was introduced into *Escherichia coli*, giving it the ability to degrade diesel fuel. The diesel-induced expression of the AlkB protein increased the diesel degradation rate from 31% to 50% after 24 h. Enhancing surfactant production is conducive to improving the accessibility of petroleum hydrocarbons to the strains, which is beneficial for improving their degradation efficiency. Wu et al. (2018) engineered *B. subtilis 168* by integrating surfactant synthesis activators, knocking out competing pathways, and enhancing the supply of fatty acid precursors, resulting in a significant increase in surfactant yield. Furthermore, different bacteria with auxiliary functions and petroleum hydrocarbon-degrading bacteria can be combined to form a consortium.

There is a way to improve the ability of the microbial consortia which is by adjusting the interspecific relationship of microbial consortia. Shuang et al. (2019) constructed a three-bacteria system with, a significantly improved degradation efficiency of phenanthrene obtained through the synergistic effect between the bacterial species. Ghorbannezhad et al. (2018) created a microbial consortium using eight fungi, three yeasts, and four bacteria, and an oil degradation assay for various combinations, including a bacterial mixed culture, a fungal mixed culture, a fungal-bacterial mixed culture, and a sequential fungal-bacterial mixed culture. The experimental results showed that the repair effect of the synergistic microbial consortia was generally significantly higher than that of a single strain. The results demonstrate that communication between different microorganisms in the microbial consortia may improve the degradation ability of petroleum hydrocarbons.

As one of the advantages of a microbial consortium, strains with additional functions can be added without affecting the degradation pathways of complex compounds. The most economically valuable strategy is to add a microbial that can use degradation products as carbon sources or substrates for the biosynthesis of high-value products. In addition, the depletion of degradation products favors the forward progression of the degradation pathway. Thus, the synthesis of high-value products increases the economic benefits of microbial consortia for degrading complex compounds (Wang et al., 2022).

Degradation of plastics by microbial consortia

A large amount of plastic is produced globally every day, but only 21% of plastics are recycled or incinerated, and most of the remainder is discarded or buried, greatly polluting the environment (Law, 2017). Under natural conditions, the whole process of plastic degradation requires a timeframe of more than 50 years (Webb et al., 2013). Plastic waste can be degraded through physical processes, chemical processes, or biodegradation (Andrady, 2011). Microbial degradation has been increasingly studied due to its safety, rapidity, and low cost. Many plastics are biodegraded by microbial consortia rather than individual strains, possibly because of the limited metabolic capacity of individual microorganisms (Qi X. et al., 2021). Yu et al. (2019) found that microorganisms in a consortium had higher biodegradation efficiency than individual strains because the potentially toxic intermediates can be removed by other microorganisms present.

Many have investigated the degradation mechanism of microbial consortia constructed by top-down approaches. Vargas-Suárez et al. (2019) selected microbial consortia from degraded foam blocks collected in landfills. They found that in the presence of microbial consortia, the carbon utilization efficiency of their strain was more efficient in degrading multiple types of complex plastics than when it was independent because of the interspecific interaction. To elucidate the mechanism by which landfill microbial consortia attack PU plastics, Gaytán et al. (2020) investigated the

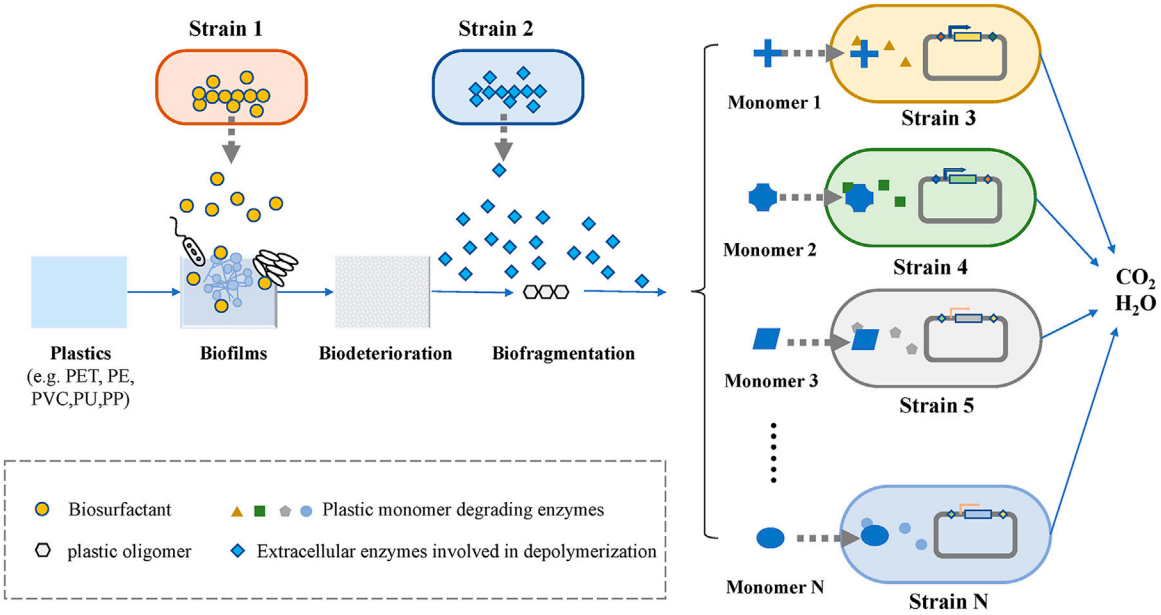


FIGURE 3
Biodegradation of plastics.

degradation of a microbial consortium selected from a municipal landfill, which was able to disperse PU in water as the sole carbon source of growth. The study showed that the degradable enzyme gene of selected microbial consortia has great potential in the direction of bioremediation. After understanding these cooperative relationships, microbial consortia can be constructed to achieve plastic degradation.

The bottom-up approach can be used to construct microbial consortia based on the different degradation pathways of plastics, whose steps are relatively clear. As shown in Figure 3, the whole process of microbial degradation can be summarized into three stages: biodeterioration, biofragmentation, and biodegradation (Zhang et al., 2022). The biodeterioration stage refers to the degradation of plastic polymer surfaces by biofilms that are formed (Ru et al., 2020). In general, biofilms of microbial consortia exhibit a better ability to degrade plastics than those of single bacteria at this stage. Due to the natural hydrophobicity of plastics, it is necessary to introduce hydrophilic functional groups on the surface of plastics to facilitate the attachment of microorganisms (Nauendorf et al., 2016). For example, the biosurfactant-producing module in the microbial consortium system for petroleum hydrocarbon degradation described above can also be used in microbial consortia for plastic degradation. Tribedi et al. (2015) demonstrated that biofilm-promoting compounds, such as mineral oil and surfactants for biofilm attachment, enhanced the biodegradation of plastics. Fungi can also play an important role in the degradation of plastics, as they can attach to plastic surfaces *via* their hyphae and provides an attachable platform for other microorganisms (Sánchez, 2020). Biofragmentation is a depolymerization step that convert plastic polymers into smaller units by the action of extracellular enzymes and free radicals (Jenkins et al., 2019). Plastic-degrading enzymes are divided into two broad categories: extracellular enzymes and intracellular enzymes. These different groups of enzymes have been found to act similarly to microbial laccases, peroxidases, lipases, esterases, and cutinases (Gan and Zhang, 2019), and they are mainly involved in depolymerizing the long carbon chains of plastic polymers to form mixtures of oligomers, dimers, and monomers. Subsequently, these monomers are then processed by different strains. Once these plastic monomers are successfully transported into cells, they undergo a series of enzymatic reactions that lead to their complete degradation into oxidative metabolites (Ho et al., 2017). The complete degradation of plastic requires the participation of a variety of enzymes, and the enzymes required for the degradation of different plastics are different. Therefore, microbial consortia are a suitable choice to increase the rate of plastics degradation, especially for a mixture of various plastics.

For artificial microbial consortia constructed by bottom-up approaches, the consortium can indirectly improve biodegradation through metabolic cross-feeding or the production of metabolites that induce co-metabolic degradation (Hu et al., 2020). The degradation pathways of plastics can be divided into different modules according to key rate-limiting enzymes and intermediates, and each module

can be assigned to different strains. These strains constitute the initial microbial consortium. This approach has the advantage of reducing the metabolic burden of each strain and increasing the tolerance of the microbial community to harsh environments through strain-to-strain interactions. Qi X. H. et al. (2021) constructed a microbial consortium consisting of *Rhodococcus*, *Pseudomonas putida,* and two engineered *B. subtilis* species for the degradation of PET. The two engineered *B. subtilis* secrete PET hydrolase and monohydroxyethyl terephthalate hydrolase to achieve the initial degradation of plastics. Then, *R. jostii* and *Pseudomonas putida* were added to degrade terephthalic acid and ethylene glycol. The final microbial consortia could completely degrade 23.2% of a PET film at room temperature. Furthermore, as a result of in-depth research on the degradation pathways of synthetic plastics, it has also become possible to build a microbial community degradation platform to degrade and convert synthetic plastics into high-value products (Qi et al., 2022; Sullivan et al., 2022).

Degradation of antibiotics and azo dyes by microbial consortia

Antibiotics have been widely used as an effective class of effective drugs, and their presence has been reported in sewage treatment plant effluent, sewage treatment plant biosolids, surface water, groundwater, and drinking water (Barancheshme and Munir, 2018; Zhang et al., 2018). Such antibiotic contamination has posed a major global threat. Some scholars claim that future bioremediation work will focus on enzymatic remediation, and biotechnology should be prioritized over chemical treatment to minimize contamination after treatment (Kumar et al., 2019). Microbial consortia showed excellent degradability in studies on the biodegradation of antibiotics. There are many types of antibiotics, and each antibiotic biodegrades in different ways. Some antibiotics are so complex that they require the cooperation of several strains to be completely degraded. Thus, microbial consortia also have advantages in degrading antibiotics.

Using the top-down approach to construct microbial consortia for the degradation of antibiotics is common at present. *Firmicutes* and *Bacteroides*, represented by *Bacillus* and *Flavobacterium*, are the main bacteria in sulfa-degrading consortia. These microbial consortia obviously can degrade sulfonamides, and almost half of the antibiotics can be degraded after 1 week, with an average degradation rate of 78.3% after 4 weeks (Liao et al., 2016b). Activated sludge is a common research object in natural microbial consortia for the degradation of antibiotics at present. Many researchers have used bacterial liquids in the activated sludge of sewage treatment plants to conduct experiments aiming to study the degradation characteristics and influencing factors of antibiotics degradation. For example, two microbial consortia isolated from activated sludge were constructed to degrade sulfamethoxazole (Larcher and Yargeau, 2011). And their degradation rates increased after sulfamethoxazole was pretreated with ozone (Larcher and Yargeau, 2013). Dominant bacteria for antibiotic degradation were screened out from the activated sludge, and a dominant microbial consortium was constructed to degrade the drugs, which included *Microbacterium sp. BR1, Rhodococcus sp. BR2, Achromobacter sp. BR3, Ralstonia sp. HR1, Ralstonia sp. HR2* and *Tsukumurella sp. HR3*. The microbial consortia degraded sulfamethoxazole with a mineralization rate of 58.0% ± 1.3% (Bouju et al., 2012).

Based on the above research results, a bottom-up approach can be used to construct artificial microbial consortia. Su et al. (2018) built a microbial consortium containing *Streptomyces sp.* and *Bacillus licheniformis* with a high degradation capacity toward β-cypermethrin, where 88.3% of β-cypermethrin could be removed within 72 h. Further, the results of the artificial microbial consortia have excellent stability and can be used in environmental restoration. Wu et al. (2020) constructed a co-culture system that could be applied to actual sewage for bioremediation, which degraded more than 80% of the tetracycline after 10 days.

Azo dyes are the most widely used synthetic dyes in textile and garment printing and dyeing. In the production and use processes of the dyes, about 10%–15% is discharged into the environment without treatment, which seriously affects the health of the contacts. These dyes have a strong solubilizing ability in water and are difficult to be removed by traditional approaches (Lellis et al., 2019). There are many approaches to treating azo dye sewage, among which microbial decolorization is considered to be the most effective and environmentally friendly. The first step in the bacterial degradation of azo dyes is to destroy the azo bonds in the dye molecules. The decolorization of azo dyes by fungi begins with hyphal adsorption, followed by the secretion of extracellular enzymes to break chemical bonds. Azo bond cleavage reactions can occur both extracellularly and intracellularly, which is favorable for the cooperation of microorganisms to degrade azo dyes. Most importantly, a single microorganism will produce toxic aromatic amines during the degradation of azo dyes, whereas microbial consortia will not (Joshi et al., 2008). This is why such a decolorization approach with microbial consortia is promoted.

However, the underlying molecular mechanism of synergistic metabolism in the microbial consortia system has not been revealed. Therefore, the current microbial consortia are mainly constructed by the top-down approach. Shanmugam et al. (2017) explored this mechanism through molecular biotechnology, finding that their microbial consortia system could biodegrade and mineralize azo dyes to a higher degree due to the synergistic relation and division of labor in the consortia. Microbial consortia have also shown excellent performance in practical applications of azo dye degradation. For example, Selim et al. (2021) isolated 21 fungi that could degrade azo dyes from

contaminated soil. Textile sewage treated with microbial consortia systems can be used to irrigate non-edible plants and alleviate the global water shortage. Although the use of azo dyes has been limited, their environmental impact is still serious. Therefore, the ability of microbial consortia to degrade azo dyes deserves further study and contributes to environmental protection stategy.

Treatment of sewage by microbial consortia

With the rapid development of modern industrialization and economic globalization, the large amount of sewage discharged by various industries is becoming a serious global environmental problem (Kong et al., 2018; Sierra et al., 2018). Traditional treatment systems are usually expensive, demand massive amounts of energy, and are often still incapable of solving all challenges associated with sewage. Using microbial consortia to treat sewage is a relatively clean and efficient approach, especially in the treatment of sewage eutrophication (Plöhn et al., 2021). The microbial consortia often studied for wastewater treatment are bacto-algae consortia developed from activated sludge. Microalgae can switch autotrophic and heterotrophic metabolism depending on the availability of carbon sources and nutrients in the surrounding environment. Therefore, microalgae are a popular candidate for building microbial consortia in water (Raja et al., 2008; Kumar et al., 2010; Subashchandrabose et al., 2013; Wijffels et al., 2013; Li et al., 2016). Microbial consortia for wastewater treatment constructed by the top-down approach have showed good stability, but with less screening process than the bottom-up approach. Therefore, they will not be discussed much here.

For microbial consortia constructed by the bottom-up approach, the combination of microalgae and bacteria showed a beneficial promoting effect. The oxygen and carbon dioxide in algae and bacteria is beneficial for growth, where algae secretions are the main carbon sources (carbohydrates, proteins, and fats) for bacteria. The metabolites of bacteria can be used as promoters for algae growth. In addition, the cell surface of microalgae can provide a stable habitat for bacteria (Ramanan et al., 2016). Bacteria break down organic matter into mineral forms and secrete extracellular metabolites such as auxin and vitamin B12, which are necessary for the growth of microalgae (Salim et al., 2014). Thus, compared with individual microorganisms, those combined with microalgae are more efficient in detoxifying organic and inorganic pollutants and removing nutrients from sewage (Subashchandrabose et al., 2013; Wijffels et al., 2013; Xiong et al., 2017). One study experimentally compared the arsenic accumulation and transformation of *Chlorella vulgaris, Aspergillus oryzae*, and bacto-algae pellets under different concentrations of arsenic and phosphorus. Among all the treatments, the removal efficiency of the bacto-algae ball was the highest and its ability to accumulate arsenic was the strongest (Li B. et al., 2019). Similarly, constructed microbial consortia have shown advantages in treating wastewater eutrophication. Mujtaba et al. (2017) studied the simultaneous removal of nutrients (ammonium and phosphate) and COD in a co-culture system of *Chlorella vulgaris* and *P. putida*. They found that the removal of nutrients and COD by the co-culture system was higher than that of each individual culture system, indicating that the nutrient absorption capacity of *Pseudomonas* putida was improved in the consortia. Thus, the combined use of microalgae and microbial consortia has broad prospects in sewage treatment.

Consolidated bioprocessing by microbial consortia

Consolidated bioprocessing (CBP) is considered one of the most potent and cost-effective ways to produce biofuels and other high-value products. It can complete the production of lignocellulose-degrading enzymes, the hydrolysis of lignocellulose, and microbial fermentation in one step. However, it is difficult to find a suitable microorganism to produce all the enzymes required for the degradation of lignocellulose and the production of high-value-added products. A promising alternative is bioprocessing based on microbial consortia. Most of the current approaches for constructing CBP microbial consortia are bottom-up approaches due to their remarkable controllability.

For the microbial consortia constructed by the bottom-up approach, some experiments verified that microbial consortia can indeed improve the efficiency of CBP. Zuroff et al. (2013) found that consortia of *C. phytofermentans* and *S. cerevisiae* produced ethanol from α-cellulose more efficiently than monocultures. If an artificial microbial consortium is constructed according to the division of labor among strains, the efficiency of CBP could be greatly improved. For example, the microbial consortia of *Zymomonas mobilis* and *Candida tropicalis* can convert enzymatically hydrolyzed lignocellulosic to ethanol with a yield reaching 97.7% (Patle and Lal, 2007). In addition to the production of ethanol, CBP can be used to produce other compounds such as halomethanes and lactic acid. Bayer et al. (2009) used a microbial consortium of engineered yeast and the cellulolytic bacterium *Actinotalea fermentans* to produce halomethanes from raw switchgrass, corn stover, bagasse, and poplar. Shahab et al. (2018) assembled an artificial microbial consortium of the cellulolytic enzyme-secreting aerobic fungus *Trichoderma reesei* with facultative anaerobic lactic acid bacteria. The results showed that the theoretically maximal lactic acid yield was obtained in the experiment.

Many experimental results show that substrate degradation efficiency is the rate-limiting step in CBP. Therefore, it is necessary to increase the substrate degradation rate to further

provide higher monosaccharide concentrations. The activity of enzymes in microbial consortia may be additionally activated and improved, thereby improving the efficiency of the CBP system (Kuhar et al., 2015). To further investigate the factors affecting the secretion of degradative enzymes in microbial consortia, Puentes-Tellez and Salles, 2018 applied a reductive screening approach based on molecular phenotype, identification, and metabolic characterization to select the desired microbial consortia. They found a minimally active microbial consortium with efficient lignocellulose-degrading ability. The degradation potential of the least active microbial consortia reached 96.5%. The enhanced degradation efficiency of lignocellulose by the mixed bacteria was more obvious in another experiment, which established a microbial consortium of *Serratia sp.* and *Arthrobacter sp.* to improve cellulose degradation. The enzymatic activity was increased by 30%–70% after co-cultivation. In addition, the degradation rate of the microbial consortia was increased by more than 30%. In another application direction, the use of microbial consortia with ligninolytic degradation ability can significantly increase the lignocellulose degradation rate in a fixed fluidized bed reactor (An et al., 2022). When the problem of the degradation efficiency of lignocellulose is solved, the construction of microbial consortia becomes much clearer. When isolated cellulose-degrading microbial consortia were co-cultured with *Clostridium acetobutyricum*, the utilization rate of cellulose was greatly improved, and a relatively high butanol product concentration was obtained (Wang et al., 2015).

Conclusion

Most current research on complex-compound-degrading microbial consortia has focused on native microbial consortia isolated from the environment. However, we believe that artificial microbial consortia are the direction of future research. With the help of metabolic engineering and synthetic biology, the construction of microbial consortia systems shows a strong degradation potential, which serves a new approach for the efficient utilization of complex substrates and the remediation of the environment. Although the mechanism of intercellular communication in large microbial consortia is still unclear, and the regulatory means are imperfect, it is predicted that with the deepening of the relevant research, the strong metabolic capacity and robustness of artificial microbial consortia will promote their use in the field of degrading complex compounds.

Author contributions

ZC: conceptualization, writing—original draft; WY: writing—review and editing; MD: conceptualization, writing—review and editing, supervision; YY: supervision. All authors have read and agreed to the published version of the manuscript.

Funding

This work was funded by the National Key Research and Development Program of China (2018YFA0902100), and National Natural Science Foundation of China (22278310).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbasian, F., Lockington, R., Mallavarapu, M., and Naidu, R. (2015). A comprehensive review of aliphatic hydrocarbon biodegradation by bacteria. *Appl. Biochem. Biotechnol.* 176 (3), 670–699. doi:10.1007/s12010-015-1603-5

Ahmad, F., Zhu, D., and Sun, J. (2020). Bacterial chemotaxis: A way forward to aromatic compounds biodegradation. *Environ. Sci. Eur.* 32 (1), 52. doi:10.1186/s12302-020-00329-2

An, X., Zong, Z., Zhang, Q., Li, Z., Zhong, M., Long, H., et al. (2022). Novel thermo-alkali-stable cellulase-producing Serratia sp. AXJ-M cooperates with Arthrobacter sp. AXJ-M1 to improve degradation of cellulose in papermaking black liquor. *J. Hazard. Mater.* 421, 126811. doi:10.1016/j.jhazmat.2021.126811

Andrady, A. L. (2011). Microplastics in the marine environment. *Mar. Pollut. Bull.* 62 (8), 1596–1605. doi:10.1016/j.marpolbul.2011.05.030

Azubuike, C. C., Chikere, C. B., and Okpokwasili, G. C. (2016). Bioremediation techniques-classification based on site of application: Principles, advantages, limitations and prospects. *World J. Microbiol. Biotechnol.* 32 (11), 180. doi:10.1007/s11274-016-2137-x

Bacosa, H. P., and Inoue, C. (2015). Polycyclic aromatic hydrocarbons (PAHs) biodegradation potential and diversity of microbial consortia enriched from tsunami sediments in Miyagi, Japan. *J. Hazard. Mater.* 283, 689–697. doi:10.1016/j.jhazmat.2014.09.068

Barancheshme, F., and Munir, M. (2018). Strategies to combat antibiotic resistance in the wastewater treatment plants. *Front. Microbiol.* 8, 2603. doi:10.3389/fmicb.2017.02603

Bardají, D. K. R., Moretto, J. A. S., Furlan, J. P. R., and Stehling, E. G. (2020). A mini-review: Current advances in polyethylene biodegradation. *World J. Microbiol. Biotechnol.* 36 (2), 32. doi:10.1007/s11274-020-2808-5

Bayer, T. S., Widmaier, D. M., Temme, K., Mirsky, E. A., Santi, D. V., and Voigt, C. A. (2009). Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* 131 (18), 6508–6515. doi:10.1021/ja809461u

Biswas, A., Mailapalli, D. R., and Raghuwanshi, N. S. (2021). Treated municipal wastewater to fulfil crop water footprints and irrigation demand – A review. *Water Supply* 21 (4), 1398–1409. doi:10.2166/ws.2021.031

Bouju, H., Ricken, B., Beffa, T., Corvini, P. F. X., and Kolvenbach, B. A. (2012). Isolation of bacterial strains capable of sulfamethoxazole mineralization from an acclimated membrane bioreactor. *Appl. Environ. Microbiol.* 78 (1), 277–279. doi:10.1128/aem.05888-11

Costa, A. S., Romao, L. P. C., Araujo, B. R., Lucas, S. C. O., Maciel, S. T. A., Wisniewski, A., Jr., et al. (2012). Environmental strategies to remove volatile aromatic fractions (BTEX) from petroleum industry wastewater using biomass. *Bioresour. Technol.* 105, 31–39. doi:10.1016/j.biortech.2011.11.096

Dhar, K., Subashchandrabose, S. R., Venkateswarlu, K., Krishnan, K., and Megharaj, M. (2020). Anaerobic microbial degradation of polycyclic aromatic hydrocarbons: A comprehensive review. *Rev. Environ. Contam. Toxicol.* 251, 25–108. doi:10.1007/398_2019_29

Embree, M., Liu, J. K., Al-Bassam, M. M., and Zengler, K. (2015). Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc. Natl. Acad. Sci* 112 (50), 15450–15455. doi:10.1073/pnas.1506034112

Friedman, J., Higgins, L. M., and Gore, J. (2017). Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* 1 (5), 0109. doi:10.1038/s41559-017-0109

Gan, Z., and Zhang, H. (2019). Pmbd: A comprehensive plastics microbial biodegradation database. *Database.* 2019, baz119. doi:10.1093/database/baz119

Garcia, I., and Oliveira, V. (2013). "Microbial hydrocarbon degradation: Efforts to understand biodegradation in petroleum reservoirs," in *Biodegradation*. Editors C. Rolando and R. Francisca (Rijeka: IntechOpen), Ch. 3.

Gaytán, I., Sánchez-Reyes, A., Burelo, M., Vargas-Suárez, M., Liachko, I., Press, M., et al. (2020). Degradation of recalcitrant polyurethane and xenobiotic additives by a selected landfill microbial community and its biodegradative potential revealed by proximity ligation-based metagenomic analysis. *Front. Microbiol.* 10, 2986. doi:10.3389/fmicb.2019.02986

Ge, C., Yu, Z., Sheng, H., Shen, X., Sun, X., Zhang, Y., et al. (2022). Redesigning regulatory components of quorum-sensing system for diverse metabolic control. *Nat. Commun.* 13 (1), 2182. doi:10.1038/s41467-022-29933-x

Ghorbannezhad, H., Moghimi, H., and Dastgheib, S. M. M. (2018). Evaluation of heavy petroleum degradation using bacterial-fungal mixed cultures. *Ecotoxicol. Environ. Saf.* 164, 434–439. doi:10.1016/j.ecoenv.2018.08.046

Ghosal, D., Ghosh, S., Dutta, T. K., and Ahn, Y. (2016). Current state of knowledge in microbial degradation of polycyclic aromatic hydrocarbons (PAHs): A review. *Front. Microbiol.* 7, 1369. doi:10.3389/fmicb.2016.01369

Goers, L., Freemont, P., and Polizzi, K. M. (2014). Co-Culture systems and technologies: Taking synthetic biology to the next level. *J. R. Soc. Interface* 11 (96), 20140065. doi:10.1098/rsif.2014.0065

Grilli, J., Barabas, G., Michalska-Smith, M. J., and Allesina, S. (2017). Higher-order interactions stabilize dynamics in competitive network models. *Nature* 548 (7666), 210–213. doi:10.1038/nature23273

Gu, H., Lou, J., Wang, H., Yang, Y., Wu, L., Wu, J., et al. (2016). Biodegradation, biosorption of phenanthrene and its trans-membrane transport by massilia sp WF1 and phanerochaete chrysosporium. *Front. Microbiol.* 7, 38. doi:10.3389/fmicb.2016.00038

Gupta, S., Pathak, B., and Fulekar, M. H. (2015). Molecular approaches for biodegradation of polycyclic aromatic hydrocarbon compounds: A review. *Rev. Environ. Sci. Biotechnol.* 14 (2), 241–269. doi:10.1007/s11157-014-9353-3

Harcombe, W. R., Chacon, J. M., Adamowicz, E. M., Chubiz, L. M., and Marx, C. J. (2018). Evolution of bidirectional costly mutualism from byproduct consumption. *Proc. Natl. Acad. Sci. U. S. A.* 115 (47), 12000–12004. doi:10.1073/pnas.1810949115

Ho, B. T., Roberts, T. K., and Lucas, S. (2017). An overview on biodegradation of polystyrene and modified polystyrene: The microbial approach. *Crit. Rev. Biotechnol.* 38 (2), 308–320. doi:10.1080/07388551.2017.1355293

Ho, B. T., Roberts, T. K., and Lucas, S. (2018). An overview on biodegradation of polystyrene and modified polystyrene: The microbial approach. *Crit. Rev. Biotechnol.* 38 (2), 308–320. doi:10.1080/07388551.2017.1355293

Holliger, C., Gaspard, S., Glod, G., Heijman, C., Schumacher, W., Schwarzenbach, R. P., et al. (1997). Contaminated environments in the subsurface and bioremediation: Organic contaminants. *FEMS Microbiol. Rev.* 20 (3-4), 517–523. doi:10.1111/j.1574-6976.1997.tb00334.x

Hu, B., Wang, M., Geng, S., Wen, L., Wu, M., Nie, Y., et al. (2020). Metabolic exchange with non-alkane-consuming Pseudomonas stutzeri slg510a3-8 improves *n*-alkane biodegradation by the alkane degrader *Dietzia* sp. strain DQ12-45-1b. *Appl. Environ. Microbiol.* 86 (8), e02931–e02919. doi:10.1128/AEM.02931-19

Ibrar, M., and Zhang, H. (2020). Construction of a hydrocarbon-degrading consortium and characterization of two new lipopeptides biosurfactants. *Sci. Total Environ.* 714, 136400. doi:10.1016/j.scitotenv.2019.136400

Jawed, K., Yazdani, S. S., and Koffas, M. A. (2019). Advances in the development and application of microbial consortia for metabolic engineering. *Metab. Eng. Commun.* 9, e00095. doi:10.1016/j.mec.2019.e00095

Jenkins, S., Quer, A. M. i., Fonseca, C., and Varrone, C. (2019). "Microbial degradation of plastics: New plastic degraders, mixed cultures and engineering strategies," in *Soil microenvironment for bioremediation and polymer production*, 213–238.

Jiang, Y., Lv, Y., Wu, R., Lu, J., Dong, W., Zhou, J., et al. (2020). Consolidated bioprocessing performance of a two-species microbial consortium for butanol production from lignocellulosic biomass. *Biotechnol. Bioeng.* 117 (10), 2985–2995. doi:10.1002/bit.27464

Johnston, T. G., Yuan, S.-F., Wagner, J. M., Yi, X., Saha, A., Smith, P., et al. (2020). Compartmentalized microbes and co-cultures in hydrogels for on-demand bioproduction and preservation. *Nat. Commun.* 11 (1), 563. doi:10.1038/s41467-020-14371-4

Jones, J. A., Vernacchio, V. R., Collins, S. M., Shirke, A. N., Xiu, Y., Englaender, J. A., et al. (2017). Complete biosynthesis of anthocyanins using *E. coli* polycultures. *Mbio* 8 (3), e00621–17. doi:10.1128/mBio.00621-17

Jones, J. A., and Wang, X. (2018). Use of bacterial co-cultures for the efficient production of chemicals. *Curr. Opin. Biotechnol.* 53, 33–38. doi:10.1016/j.copbio.2017.11.012

Joshi, T., Iyengar, L., Singh, K., and Garg, S. (2008). Isolation, identification and application of novel bacterial consortium TJ-1 for the decolourization of structurally different azo dyes. *Bioresour. Technol.* 99 (15), 7115–7121. doi:10.1016/j.biortech.2007.12.074

Kaeberlein, T., Lewis, K., and Epstein, S. S. (2002). Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* 296 (5570), 1127–1129. doi:10.1126/science.1070633

Kelsic, E. D., Zhao, J., Vetsigian, K., and Kishony, R. (2015). Counteraction of antibiotic production and degradation stabilizes microbial communities. *Nature* 521 (7553), 516–519. doi:10.1038/nature14485

Kong, Z., Li, L., Kurihara, R., Kubota, K., and Li, Y.-Y. (2018). Anaerobic treatment of N, N-dimethylformamide-containing wastewater by co-culturing two sources of inoculum. *Water Res.* 139, 228–239. doi:10.1016/j.watres.2018.03.078

Kuhar, F., Castiglia, V., and Levin, L. (2015). Enhancement of laccase production and malachite green decolorization by co-culturing Ganoderma lucidum and *Trametes versicolor* in solid-state fermentation. *Int. Biodeterior. Biodegrad.* 104, 238–243. doi:10.1016/j.ibiod.2015.06.017

Kumar, A., Ergas, S., Yuan, X., Sahu, A., Zhang, Q., Dewulf, J., et al. (2010). Enhanced CO2 fixation and biofuel production via microalgae: Recent developments and future directions. *Trends Biotechnol.* 28 (7), 371–380. doi:10.1016/j.tibtech.2010.04.004

Kumar, M., Jaiswal, S., Sodhi, K. K., Shree, P., Singh, D. K., Agrawal, P. K., et al. (2019). Antibiotics bioremediation: Perspectives on its ecotoxicity and resistance. *Environ. Int.* 124, 448–461. doi:10.1016/j.envint.2018.12.065

Larcher, S., and Yargeau, V. (2011). Biodegradation of sulfamethoxazole by individual and mixed bacteria. *Appl. Microbiol. Biotechnol.* 91 (1), 211–218. doi:10.1007/s00253-011-3257-8

Larcher, S., and Yargeau, V. (2013). The effect of ozone on the biodegradation of 17 alpha-ethinylestradiol and sulfamethoxazole by mixed bacterial cultures. *Appl. Microbiol. Biotechnol.* 97 (5), 2201–2210. doi:10.1007/s00253-012-4054-8

Law, K. L. (2017). Plastics in the marine environment. *Ann. Rev. Mar. Sci.* 9 (1), 205–229. doi:10.1146/annurev-marine-010816-060409

Lawson, C. E., Harcombe, W. R., Hatzenpichler, R., Lindemann, S. R., Löffler, F. E., O'Malley, M. A., et al. (2019). Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* 17 (12), 725–741. doi:10.1038/s41579-019-0255-9

Lellis, B., Fávaro-Polonio, C., Pamphile, J. A., and Polonio, J. C. (2019). Effects of textile dyes on health and the environment and bioremediation potential of living organisms. *Biotechnol. Res. Innovation* 3 (2), 275–290. doi:10.1016/j.biori.2019.09.001

Li, B., Zhang, T., and Yang, Z. (2019a). Immobilizing unicellular microalga on pellet-forming filamentous fungus: Can this provide new insights into the remediation of arsenic from contaminated water? *Bioresour. Technol.* 284, 231–239. doi:10.1016/j.biortech.2019.03.128

Li, C., Xiao, S., and Ju, L.-K. (2016). Cultivation of phagotrophic algae with waste activated sludge as a fast approach to reclaim waste organics. *Water Res.* 91, 195–202. doi:10.1016/j.watres.2016.01.021

Li, H., and Ding, M. (2021). Advances in biodegradation of petroleum hydrocarbons. *Sheng Wu Gong Cheng Xue Bao* 37 (8), 2765–2778. doi:10.13345/j.cjb.200611

Li, H. L., Lai, R. Q., Jin, Y. L., Fang, X. X., Cui, K., Sun, S. S., et al. (2020). Directional culture of petroleum hydrocarbon degrading bacteria for enhancing crude oil recovery. *J. Hazard. Mater.* 390, 122160. doi:10.1016/j.jhazmat.2020.122160

Li, X., Zhou, Z., Li, W., Yan, Y., Shen, X., Wang, J., et al. (2022). Design of stable and self-regulated microbial consortia for chemical synthesis. *Nat. Commun.* 13 (1), 1554. doi:10.1038/s41467-022-29215-6

Li, Z., Wang, X., and Zhang, H. (2019b). Balancing the non-linear rosmarinic acid biosynthetic pathway by modular co-culture engineering. *Metab. Eng.* 54, 1–11. doi:10.1016/j.ymben.2019.03.002

Liao, X., Li, B., Zou, R., Dai, Y., Xie, S., and Yuan, B. (2016a). Biodegradation of antibiotic ciprofloxacin: Pathways, influential factors, and bacterial community structure. *Environ. Sci. Pollut. Res.* 23 (8), 7911–7918. doi:10.1007/s11356-016-6054-1

Liao, X., Li, B., Zou, R., Xie, S., and Yuan, B. (2016b). Antibiotic sulfanilamide biodegradation by acclimated microbial populations. *Appl. Microbiol. Biotechnol.* 100 (5), 2439–2447. doi:10.1007/s00253-015-7133-9

Lu, H., Villada, J. C., and Lee, P. K. H. (2019). Modular metabolic engineering for biobased chemical production. *Trends Biotechnol.* 37 (2), 152–166. doi:10.1016/j.tibtech.2018.07.003

Luo, Q., Ying, H., Hou, D. Y., Zhang, J. G., and Shen, X. R. (2015a). GPo1 <italic>alkB</italic> gene expression for improvement of the degradation of diesel oil by a bacterial consortium. *Braz. J. Microbiol.* 46 (3), 649–657. doi:10.1590/S1517-838246320120226

Luo, X., Tsao, C. Y., Wu, H. C., Quan, D. N., Payne, G. F., Rubloff, G. W., et al. (2015b). Distal modulation of bacterial cell-cell signalling in a synthetic ecosystem using partitioned microfluidics. *Lab. Chip* 15 (8), 1842–1851. doi:10.1039/c5lc00107b

Malik, K., Salama, E.-S., El-Dalatony, M. M., Jalalah, M., Harraz, F. A., Al-Assiri, M. S., et al. (2021). Co-fermentation of immobilized yeasts boosted bioethanol production from pretreated cotton stalk lignocellulosic biomass: Long-term investigation. *Industrial Crops Prod.* 159, 113122. doi:10.1016/j.indcrop.2020.113122

Mashruwala, A. A., Qin, B., and Bassler, B. L. (2022). Quorum-sensing- and type VI secretion-mediated spatiotemporal cell death drives genetic diversity in *Vibrio cholerae*. *Cell* 185, 3966–3979. e13. doi:10.1016/j.cell.2022.09.003

Mayfield, M. M., and Stouffer, D. B. (2017). Higher-order interactions capture unexplained complexity in diverse communities. *Nat. Ecol. Evol.* 1 (3), 0062. doi:10.1038/s41559-016-0062

McCarty, N. S., and Ledesma-Amaro, R. (2019). Synthetic biology tools to engineer microbial communities for biotechnology. *Trends Biotechnol.* 37 (2), 181–197. doi:10.1016/j.tibtech.2018.11.002

McDonald, I. R., Miguez, C. B., Rogge, G., Bourque, D., Wendlandt, K. D., Groleau, D., et al. (2006). Diversity of soluble methane monooxygenase-containing methanotrophs isolated from polluted environments. *FEMS Microbiol. Lett.* 255 (2), 225–232. doi:10.1111/j.1574-6968.2005.00090.x

Mujtaba, G., Rizwan, M., and Lee, K. (2017). Removal of nutrients and COD from wastewater using symbiotic co-culture of bacterium Pseudomonas putida and immobilized microalga Chlorella vulgaris. *J. Industrial Eng. Chem.* 49, 145–151. doi:10.1016/j.jiec.2017.01.021

Nauendorf, A., Krause, S., Bigalke, N. K., Gorb, E. V., Gorb, S. N., Haeckel, M., et al. (2016). Microbial colonization and degradation of polyethylene and biodegradable plastic bags in temperate fine-grained organic-rich marine sediments. *Mar. Pollut. Bull.* 103 (1), 168–178. doi:10.1016/j.marpolbul.2015.12.024

Orth, J. D., Thiele, I., and Palsson, B. (2010). What is flux balance analysis? *Nat. Biotechnol.* 28 (3), 245–248. doi:10.1038/nbt.1614

Park, H., Patel, A., Hunt, K. A., Henson, M. A., and Carlson, R. P. (2020). Artificial consortium demonstrates emergent properties of enhanced cellulosic-sugar degradation and biofuel synthesis. *npj Biofilms Microbiomes* 6 (1), 59. doi:10.1038/s41522-020-00170-8

Patle, S., and Lal, B. (2007). Ethanol production from hydrolysed agricultural wastes using mixed culture of Zymomonas mobilis and *Candida tropicalis*. *Biotechnol. Lett.* 29 (12), 1839–1843. doi:10.1007/s10529-007-9493-4

Plöhn, M., Spain, O., Sirin, S., Silva, M., Escudero-Oñate, C., Ferrando-Climent, L., et al. (2021). Wastewater treatment by microalgae. *Physiol. Plant.* 173 (2), 568–578. doi:10.1111/ppl.13427

Puentes-Tellez, P. E., and Salles, J. F. (2018). Construction of effective minimal active microbial consortia for lignocellulose degradation. *Microb. Ecol.* 76 (2), 419–429. doi:10.1007/s00248-017-1141-5

Pütz, E., Gazanis, A., Keltsch, N. G., Jegel, O., Pfitzner, F., Heermann, R., et al. (2022). Communication breakdown: Into the molecular mechanism of biofilm inhibition by CeO2 nanocrystal enzyme mimics and how it can Be exploited. *ACS Nano* 16, 16091–16108. doi:10.1021/acsnano.2c04377

Qi, X. H., Ma, Y., Chang, H. C., Li, B. Z., Ding, M. Z., and Yuan, Y. J. (2021b). Evaluation of PET degradation using artificial microbial consortia. *Front. Microbiol.* 12, 778828. doi:10.3389/fmicb.2021.778828

Qi, X., Ma, Y., Chang, H., Li, B., Ding, M., and Yuan, Y. (2021a). Evaluation of PET degradation using artificial microbial consortia. *Front. Microbiol.* 12, 778828. doi:10.3389/fmicb.2021.778828

Qi, X., Yan, W., Cao, Z., Ding, M. Z., and Yuan, Y. J. (2022). Current advances in the biodegradation and bioconversion of polyethylene terephthalate. *Microorganisms* 10 (1), 39. doi:10.3390/microorganisms10010039

Qian, X., Chen, L., Sui, Y., Chen, C., Zhang, W., Zhou, J., et al. (2020). Biotechnological potential and applications of microbial consortia. *Biotechnol. Adv.* 40, 107500. doi:10.1016/j.biotechadv.2019.107500

Qian, Y., Xu, M., Deng, T., Hu, W., He, Z., Yang, X., et al. (2021). Synergistic interactions of Desulfovibrio and Petrimonas for sulfate-reduction coupling polycyclic aromatic hydrocarbon degradation. *J. Hazard. Mater.* 407, 124385. doi:10.1016/j.jhazmat.2020.124385

Raja, R., Hemaiswarya, S., Kumar, N. A., Sridhar, S., and Rengasamy, R. (2008). A perspective on the biotechnological potential of microalgae. *Crit. Rev. Microbiol.* 34 (2), 77–88. doi:10.1080/10408410802086783

Ramanan, R., Kim, B.-H., Cho, D.-H., Oh, H.-M., and Kim, H.-S. (2016). Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol. Adv.* 34 (1), 14–29. doi:10.1016/j.biotechadv.2015.12.003

Ramsay, J. P., Bastholm, T. R., Callum, J. V., Dinah, D. T., John, T. S., Liam, K. H., et al. (2022). An epigenetic switch activates bacterial quorum sensing and horizontal transfer of an integrative and conjugative element. *Nucleic Acids Res.* 50 (2), 975–988. doi:10.1093/nar/gkab1217

Roell, G. W., Zha, J., Carr, R. R., Koffas, M. A., Fong, S. S., and Tang, Y. J. (2019). Engineering microbial consortia by division of labor. *Microb. Cell Fact.* 18 (1), 35. doi:10.1186/s12934-019-1083-3

Ron, E. Z., and Rosenberg, E. (2014). Enhanced bioremediation of oil spills in the sea. *Curr. Opin. Biotechnol.* 27, 191–194. doi:10.1016/j.copbio.2014.02.004

Ru, J., Huo, Y., and Yang, Y. (2020). Microbial degradation and valorization of plastic wastes. *Front. Microbiol.* 11, 442. doi:10.3389/fmicb.2020.00442

Said, S. B., and Or, D. (2017). Synthetic microbial ecology: Engineering habitats for modular consortia. *Front. Microbiol.* 8, 1125. doi:10.3389/fmicb.2017.01125

Sajna, K. V., Sukumaran, R. K., Gottumukkala, L. D., and Pandey, A. (2015). Crude oil biodegradation aided by biosurfactants from Pseudozyma sp NII 08165 or its culture broth. *Bioresour. Technol.* 191, 133–139. doi:10.1016/j.biortech.2015.04.126

Salim, S., Kosterink, N. R., Wacka, N. D. T., Vermue, M. H., and Wijffels, R. H. (2014). Mechanism behind autoflocculation of unicellular green micro algae Ettlia texensis. *J. Biotechnol.* 174, 34–38. doi:10.1016/j.jbiotec.2014.01.026

Samir Ali, S., Al-Tohamy, R., Khalil, M. A., Ho, S.-H., Fu, Y., and Sun, J. (2022). Exploring the potential of a newly constructed manganese peroxidase-producing yeast consortium for tolerating lignin degradation inhibitors while simultaneously decolorizing and detoxifying textile azo dye wastewater. *Bioresour. Technol.* 351, 126861. doi:10.1016/j.biortech.2022.126861

Sánchez, C. (2020). Fungal potential for the degradation of petroleum-based polymers: An overview of macro- and microplastics biodegradation. *Biotechnol. Adv.* 40, 107501. doi:10.1016/j.biotechadv.2019.107501

Selim, M. T., Salem, S. S., Mohamed, A. A., El-Gamal, M. S., Awad, M. F., and Fouda, A. (2021). Biological treatment of real textile effluent using Aspergillus flavus and Fusarium oxysporium and their consortium along with the evaluation of their phytotoxicity. *J. Fungi (Basel).* 7 (3), 193. doi:10.3390/jof7030193

Sepehri, A., Sarrafzadeh, M.-H., and Avateffazeli, M. (2020). Interaction between Chlorella vulgaris and nitrifying-enriched activated sludge in the treatment of wastewater with low C/N ratio. *J. Clean. Prod.* 247, 119164. doi:10.1016/j.jclepro.2019.119164

Shahab, R. L., Brethauer, S., Davey, M. P., Smith, A. G., Vignolini, S., Luterbacher, J. S., et al. (2020). A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose. *Science* 369 (6507), eabb1214. doi:10.1126/science.abb1214

Shahab, R. L., Luterbacher, J. S., Brethauer, S., and Studer, M. H. (2018). Consolidated bioprocessing of lignocellulosic biomass to lactic acid by a synthetic fungal-bacterial consortium. *Biotechnol. Bioeng.* 115 (5), 1207–1215. doi:10.1002/bit.26541

Shanmugam, B. K., Easwaran, S. N., Lakra, R., Deepa, P. R., and Mahadevan, S. (2017). Metabolic pathway and role of individual species in the bacterial consortium for biodegradation of azo dye: A biocalorimetric investigation. *Chemosphere* 188, 81–89. doi:10.1016/j.chemosphere.2017.08.138

Shen, Y.-P., Niu, F.-X., Yan, Z.-B., Fong, L. S., Huang, Y.-B., and Liu, J.-Z. (2020). Recent advances in metabolically engineered microorganisms for the production of aromatic chemicals derived from aromatic amino acids. *Front. Bioeng. Biotechnol.* 8, 407. doi:10.3389/fbioe.2020.00407

Shi, Y., Yang, Z., Xing, L., Zhang, X., Li, X., and Zhang, D. (2021). Recent advances in the biodegradation of azo dyes. *World J. Microbiol. Biotechnol.* 37 (8), 137. doi:10.1007/s11274-021-03110-6

Shuang, W. U., Liu, C., Yang, L., Long, H., Xuan, J., and Jiang, M. (2019). Screening, identification and degradation characteristics of threephenanthrene-degrading bacteria isolated form mangrove Soil. *Environ. Sci. Technol.* 42 (3), 73–79. doi:10.19672/j.cnki.1003-6504.2019.03.010

Sierra, J. D. M., Wang, W., Cerqueda-Garcia, D., Oosterkamp, M. J., Spanjers, H., and van Lier, J. B. (2018). Temperature susceptibility of a mesophilic anaerobic membrane bioreactor treating saline phenol-containing wastewater. *Chemosphere* 213, 92–102. doi:10.1016/j.chemosphere.2018.09.023

Song, H., Ding, M.-Z., Jia, X.-Q., Ma, Q., and Yuan, Y.-J. (2014). Synthetic microbial consortia: From systematic analysis to construction and applications. *Chem. Soc. Rev.* 43 (20), 6954–6981. doi:10.1039/C4CS00114A

Striednig, B., and Hilbi, H. (2022). Bacterial quorum sensing and phenotypic heterogeneity: How the collective shapes the individual. *Trends Microbiol.* 30 (4), 379–389. doi:10.1016/j.tim.2021.09.001

Su, H., Chen, Q., Zhao, J., Chi, Y., Jia, D., and Yao, K. (2018). Appropriate conditions of beta-cypermethrin degradation by a co-culture of microorganism. *Food Ferment. Industries* 44 (7), 8–12. doi:10.1073/pnas.80.1.178

Subashchandrabose, S. R., Ramakrishnan, B., Megharaj, M., Venkateswarlu, K., and Naidu, R. (2013). Mixotrophic cyanobacteria and microalgae as distinctive biological agents for organic pollutant degradation. *Environ. Int.* 51, 59–72. doi:10.1016/j.envint.2012.10.007

Sullivan, K. P., Werner, A. Z., Ramirez, K. J., Ellis, L. D., Bussard, J. R., Black, B. A., et al. (2022). Mixed plastics waste valorization through tandem chemical oxidation and biological funneling. *Science* 378 (6616), 207–211. doi:10.1126/science.abo4626

Tondro, H., Musivand, S., Zilouei, H., Bazarganipour, M., and Zargoosh, K. (2020). Biological production of hydrogen and acetone- butanol-ethanol from sugarcane bagasse and rice straw using co-culture of Enterobacter aerogenes and Clostridium acetobutylicum. *Biomass Bioenergy* 142, 105818. doi:10.1016/j.biombioe.2020.105818

Tribedi, P., Gupta, A. D., and Sil, A. K. (2015). Adaptation of Pseudomonas sp. AKS2 in biofilm on low-density polyethylene surface: An effective strategy for efficient survival and polymer degradation. *Bioresour. Bioprocess.* 2 (1), 14. doi:10.1186/s40643-015-0044-x

Vargas-Suárez, M., Fernández-Cruz, V., and Loza-Tavera, H. (2019). Biodegradation of polyacrylic and polyester polyurethane coatings by enriched microbial communities. *Appl. Microbiol. Biotechnol.* 103 (7), 3225–3236. doi:10.1007/s00253-019-09660-y

Varjani, S. J., Rana, D. P., Bateja, S., and Upasani, V. N. (2013). Isolation and screening for hydrocarbon utilizing bacteria (HUB) from petroleum samples. *Int. J. Curr. Microbiol. Appl. Sci.* 2 (4), 48–60.

Varjani, S. J., Rana, D. P., Jain, A. K., Bateja, S., and Upasani, V. N. (2015). Synergistic *ex-situ* biodegradation of crude oil by halotolerant bacterial consortium of indigenous strains isolated from on shore sites of Gujarat, India. *Int. Biodeterior. Biodegrad.* 103, 116–124. doi:10.1016/j.ibiod.2015.03.030

Varjani, S. J., and Upasani, V. N. (2016). Carbon spectrum utilization by an indigenous strain of *Pseudomonas aeruginosa* NCIM 5514: Production, characterization and surface active properties of biosurfactant. *Bioresour. Technol.* 221, 510–516. doi:10.1016/j.biortech.2016.09.080

Varjani, S. J., and Upasani, V. N. (2013). Comparative studies on bacterial consortia for hydrocarbon degradation. *Int. J. Innov. Res. Sci. Eng. Technol.* 2 (10), 5377–5383.

Wang, C., Li, Y., Tan, H., Zhang, A., Xie, Y., Wu, B., et al. (2019a). A novel microbe consortium, nano-visible light photocatalyst and microcapsule system to degrade PAHs. *Chem. Eng. J.* 359, 1065–1074. doi:10.1016/j.cej.2018.11.077

Wang, S., Tang, H., Peng, F., Yu, X., Su, H., Xu, P., et al. (2019b). Metabolite-based mutualism enhances hydrogen production in a two-species microbial consortium. *Commun. Biol.* 2, 82. doi:10.1038/s42003-019-0331-8

Wang, Y., Li, H., Liu, Y., Zhou, M., Ding, M., and Yuan, Y. (2022). Construction of synthetic microbial consortia for 2-keto-L-gulonic acid biosynthesis. *Synthetic Syst. Biotechnol.* 7 (1), 481–489. doi:10.1016/j.synbio.2021.12.001

Wang, Z., Cao, G., Zheng, J., Fu, D., Song, J., Zhang, J., et al. (2015). Developing a mesophilic co-culture for direct conversion of cellulose to butanol in consolidated bioprocess. *Biotechnol. Biofuels* 8, 84. doi:10.1186/s13068-015-0266-3

Webb, H. K., Arnott, J., Crawford, R. J., and Ivanova, E. P. (2013). Plastic degradation and its environmental implications with special reference to poly(ethylene terephthalate). *Polymers* 5 (1), 1–18. doi:10.3390/polym5010001

Wen, Z., Ledesma-Amaro, R., Lu, M., Jiang, Y., Gao, S., Jin, M., et al. (2020). Combined evolutionary engineering and genetic manipulation improve low pH tolerance and butanol production in a synthetic microbial Clostridium community. *Biotechnol. Bioeng.* 117 (7), 2008–2022. doi:10.1002/bit.27333

Wijffels, R. H., Kruse, O., and Hellingwerf, K. J. (2013). Potential of industrial biotechnology with cyanobacteria and eukaryotic microalgae. *Curr. Opin. Biotechnol.* 24 (3), 405–413. doi:10.1016/j.copbio.2013.04.004

Wu, Q., Zhi, Y., and Xu, Y. (2018). Systematically engineering the biosynthesis of a green biosurfactant surfactin by Bacillus subtilis 168. *Metab. Eng.* 52, 87–97. doi:10.1016/j.ymben.2018.11.004

Wu, X., Zhou, X., Wu, X., Luo, K., Gu, Y., Zhou, H., et al. (2020). Construction of tetracycline-degrading bacterial Co-culture system and community analysis of wastewater remediation. *Biotechnol. Bull.* 36 (10), 116–126. doi:10.1016/0379-0738(83)90103-2

Xiong, J.-Q., Miracle, M. B., and Jeon, B.-H. (2017). Ecotoxicological effects of enrofloxacin and its removal by monoculture of microalgal species and their consortium. *Environ. Pollut.* 226, 486–493. doi:10.1016/j.envpol.2017.04.044

Yan, R., Wang, Y., Li, J., Wang, X., and Wang, Y. (2022). Determination of the lower limits of antibiotic biodegradation and the fate of antibiotic resistant genes in activated sludge: Both nitrifying bacteria and heterotrophic bacteria matter. *J. Hazard. Mater.* 425, 127764. doi:10.1016/j.jhazmat.2021.127764

Yu, K., Yi, S., Li, B., Guo, F., Peng, X., Wang, Z., et al. (2019). An integrated meta-omics approach reveals substrates involved in synergistic interactions in a bisphenol A (BPA)-degrading microbial community. *Microbiome* 7 (1), 16. doi:10.1186/s40168-019-0634-5

Zanaroli, G., Di Toro, S., Todaro, D., Varese, G. C., Bertolotto, A., and Fava, F. (2010). Characterization of two diesel fuel degrading microbial consortia enriched from a non acclimated, complex source of microorganisms. *Microb. Cell Fact.* 9, 10. doi:10.1186/1475-2859-9-10

Zhang, C., and Hong, K. (2020). Production of terpenoids by synthetic biology approaches. *Front. Bioeng. Biotechnol.* 8, 347. doi:10.3389/fbioe.2020.00347

Zhang, H., Jia, Y., Khanal, S. K., Lu, H., Fang, H., and Zhao, Q. (2018). Understanding the role of extracellular polymeric substances on ciprofloxacin adsorption in aerobic sludge, anaerobic sludge, and sulfate-reducing bacteria sludge systems. *Environ. Sci. Technol.* 52 (11), 6476–6486. doi:10.1021/acs.est.8b00568

Zhang, H., and Stephanopoulos, G. (2016). Co-culture engineering for microbial biosynthesis of 3-amino-benzoic acid in *Escherichia coli*. *Biotechnol. J.* 11 (7), 981–987. doi:10.1002/biot.201600013

Zhang, N., Ding, M., and Yuan, Y. (2022). Current advances in biodegradation of polyolefins. *Microorganisms* 10 (8), 1537. doi:10.3390/microorganisms10081537

Zhang, Y., Shi, K., Cui, H., Han, J., Wang, H., Ma, X., et al. (2023). Efficient biodegradation of acetoacetanilide in hypersaline wastewater with a synthetic halotolerant bacterial consortium. *J. Hazard. Mater.* 441, 129926. in press. doi:10.1016/j.jhazmat.2022.129926

Zhou, K., Qiao, K., Edgar, S., and Stephanopoulos, G. (2015). Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* 33 (4), 377–383. doi:10.1038/nbt.3095

Zuroff, T. R., Xiques, S. B., and Curtis, W. R. (2013). Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic Clostridium phytofermentans/yeast co-culture. *Biotechnol. Biofuels* 6 (1), 59. doi:10.1186/1754-6834-6-59

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Brief Research Report
PUBLISHED 15 December 2022
DOI 10.3389/fbioe.2022.1098021

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Yaoping Zhang,
University of Wisconsin-Madison, United States
Zhengming Zhu,
Nanjing Tech University, China

*CORRESPONDENCE
Xia Wang,
xxwang@hubu.edu.cn
Shihui Yang,
shhyoung@hotmail.com

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 14 November 2022
ACCEPTED 01 December 2022
PUBLISHED 15 December 2022

CITATION
Tang Y, Wang Y, Yang Q, Zhang Y, Wu Y, Yang Y, Mei M, He M, Wang X and Yang S (2022), Molecular mechanism of enhanced ethanol tolerance associated with *hfq* overexpression in *Zymomonas mobilis*. *Front. Bioeng. Biotechnol.* 10:1098021. doi: 10.3389/fbioe.2022.1098021

Molecular mechanism of enhanced ethanol tolerance associated with *hfq* overexpression in *Zymomonas mobilis*

Ying Tang[1], Yi Wang[1], Qing Yang[1], Youpeng Zhang[1], Yalun Wu[1], Yongfu Yang[1], Meng Mei[1], Mingxiong He[2], Xia Wang[1]* and Shihui Yang[1]*

[1]State Key Laboratory of Biocatalysis and Enzyme Engineering, Environmental Microbial Technology Center of Hubei Province and School of Life Sciences, Hubei University, Wuhan, China, [2]Key Laboratory of Development and Application of Rural Renewable Energy, Biomass Energy Technology Research Centre, Biogas Institute of Ministry of Agriculture, Ministry of Agriculture, Chengdu, China

Zymomonas mobilis is a promising microorganism for industrial bioethanol production. However, ethanol produced during fermentation is toxic to *Z. mobilis* and affects its growth and bioethanol production. Although several reports demonstrated that the RNA-binding protein Hfq in *Z. mobilis* contributes to the tolerance against multiple lignocellulosic hydrolysate inhibitors, the role of Hfq on ethanol tolerance has not been investigated. In this study, *hfq* in *Z. mobilis* was either deleted or overexpressed and their effects on cell growth and ethanol tolerance were examined. Our results demonstrated that *hfq* overexpression improved ethanol tolerance of *Z. mobilis,* which is probably due to energy saving by downregulating flagellar biosynthesis and heat stress response proteins, as well as reducing the reactive oxygen species induced by ethanol stress *via* upregulating the sulfate assimilation and cysteine biosynthesis. To explore proteins potentially interacted with Hfq, the TEV protease mediated Yeast Endoplasmic Reticulum Sequestration Screening system (YESS) was established in *Z. mobilis*. YESS results suggested that Hfq may modulate the cytoplasmic heat shock response by interacting with the heat shock proteins DnaK and DnaJ to deal with the ethanol inhibition. This study thus not only revealed the underlying mechanism of enhanced ethanol tolerance by *hfq* overexpression, but also provided an alternative approach to investigate protein-protein interactions in *Z. mobilis*.

KEYWORDS

***Zymomonas mobilis*, ethanol tolerance, Hfq, ROS-reactive oxygen species, protein-protein interaction (PPI), sulfate assimilation, cysteine biosynthesis, yeast endoplasmic reticulum sequestration screening system (YESS)**

Introduction

With the increasing demand for unsustainable fossil fuel reserves, the production of biofuels from renewable resources has become increasingly important and attracted considerable attentions such as bioethanol production by a variety of microorganisms, such as *Saccharomyces cerevisiae*, *Escherichia coli*, and *Zymomonas mobilis* by fermenting renewable resources such as lignocellulosic biomass.

Z. mobilis is a natural ethanologen exhibiting advantages of high sugar uptake, high specific ethanol productivity and yield, as well as high ethanol tolerance up to 16% (v/v) in batch fermentation (Rogers et al., 2007; Yang et al., 2013; Yang et al., 2016). Moreover, it does not require controlled oxygen addition during the fermentation process (He et al., 2014; Zhang K. et al., 2019). In addition, *Z. mobilis* has been genetically modified to enhance its robustness against different stresses including lignocellulosic hydrolysate inhibitors such as acetate, vanillin, and furfural (Yang Y. et al., 2018; Wang et al., 2018; Zhang K. et al., 2019). These advantages make *Z. mobilis* to be a promising microbial cell factory for industrial lignocellulosic bioethanol production.

However, the accumulation of ethanol produced during fermentation is still toxic to *Z. mobilis*, which is a bottleneck for bioethanol production improvement. Ethanol is a chaotropic compound that can promote changes in membrane composition and influence the structure and function of macromolecules such as proteins, nucleic acids, and lipids. At high concentrations, ethanol impedes the specific cell growth rate and viability of *Z. mobilis* cells and ultimately results in the death of the microorganism (Thanonkeo et al., 2007; Tan et al., 2016). To better understand and address these limitations, it is essential to obtain *Z. mobilis* mutant strains with improved ethanol tolerance.

Previous studies on molecular response to ethanol stress in *Z. mobilis* by transcriptomics and proteomics approaches demonstrated that multiple genes involved in different cellular processes were differentially regulated against ethanol stress such as membrane biogenesis, respiratory chain, DNA replication and recombination, transcriptional regulation, and general stress responses (He et al., 2012; Yang et al., 2013). Recently, Pallach *et al.* reported that one of the two cell surface exopolysaccharides in *Z. mobilis*, the galactose containing polymer (PS1), may have a crucial role in ethanol tolerance (Pallach et al., 2018). Another study in two ethanol-tolerant mutants exhibited that *clpP*, *spoT*/*relA*, and *clpB* genes seem to contribute to the ethanol tolerance in *Z. mobilis* (Carreon-Rodriguez et al., 2019).

Several global regulators including transcription factors were related to ethanol stress. For example, the heterologous expression of a global regulator gene *irrE* from *Deinococcus radiodurans* conferred high ethanol tolerance to *Z. mobilis* (Zhang et al., 2010). The random mutagenesis of gene encoding the native global transcription sigma factor (σ^{70}, RpoD) *via* global transcription machinery engineering (gTME) enhanced ethanol resistance through modulating the transcriptional level (Tan et al., 2016). In addition, many efforts have also been focused on distinguishing the transcriptional regulators and transcription machinery, including small RNAs and RNA chaperones.

Hfq is a ubiquitous conserved Sm-like RNA-binding protein, which was originally identified as a host factor required for bacteriophage Qβ RNA replication in *E. coli*. Like eukaryotic and archaeal Sm/Lsm proteins, Hfq plays several roles in bacterial RNA metabolism, particularly to stabilize sRNAs (small non-coding RNAs) and promote their interactions with mRNAs leading to modulating stability and/or translation of these targets (Vogel and Luisi, 2011; Updegrove et al., 2016; Dos Santos et al., 2019). Beyond its sRNA-mediated regulation, Hfq was recently found to bind rRNA acting as a new ribosome biogenesis factor, and to bind tRNAs involved in the accuracy of protein synthesis. Moreover, Hfq is also capable to establish many protein-protein interactions (Yonekura et al., 2013; Dos Santos et al., 2019).

Given the central role played by Hfq in sRNA-mediated gene regulation in many bacteria, the protein has been widely recognized as a pleiotropic regulator of cell physiology, which particularly affects the cell response to multiple stresses. For example, at least three sRNAs, DsrA, RprA, and ArcZ interact with Hfq to positively regulate the *rpoS* transcript, which encodes the stress response sigma factor σ^{S} and controls >10% of all protein-coding genes in *E. coli* (Soper and Woodson, 2008; Frohlich and Gottesman, 2018).

Hfq in *Z. mobilis*, encoded by *ZMO0347*, was also discovered to contribute to multiple stress responses such as lignocellulosic hydrolysate inhibitor tolerance (Yang et al., 2010). Overexpression of *hfq* enabled the recombinant *Z. mobilis* an improved resistance to acetic acid, furfural and sugarcane bagasse hydrolysate compared to the parental strain (Nouri et al., 2020). Moreover, the 5′ untranslated regions (5′ UTRs) of *hfq* gene as the regulatory element was revealed to downregulate downstream gene expression under ethanol stress in *Z. mobilis* (Cho et al., 2017). These findings suggested an important role of Hfq in dealing with ethanol stress in *Z. mobilis*. However, the full repertoire of Hfq-dependent gene regulation response to ethanol has not been elucidated, although omics efforts have been carried out to understand ethanol stress responses in *Z. mobilis* (Zhang et al., 2010; Tan et al., 2016; Han et al., 2020). In this study, the role of *hfq* on ethanol tolerance in *Z. mobilis* was investigated by constructing mutant strains of *hfq* deletion or overexpression, and their molecular responses were characterized using genetic approaches and transcriptomic study.

Materials and methods

Strains, media, and growth conditions

Bacterial strains and plasmids used in this study are listed in Supplementary Table S1. *Z. mobilis* ZM4 (ATCC 31821) was used as the parental strain in this study. Generally, ZM4 and its derivative strains are cultured at 30°C with shaking at 100 rpm in RM medium (50 g/L glucose, 10 g/L yeast extract, 2 g/L KH_2PO_4, and 1.5% agar for solid). To avoid the effect of other amino acid nutrients in RM medium, MM medium (50 g/L glucose, 1 g/L KH_2PO_4, 1 g/L K_2HPO_4, 1 g/L $(NH_4)_2SO_4$, 0.5 g/L NaCl, 0.42 g/L $MgCl_2$ $6H_2O$, 0.001 g/L calcium pantothenate, and 1.5% agar for solid) was used in this study for Na_2SO_4 supplementation experiments, and 4 g/L Na_2SO_4 was added. The final concentration of exogenous ethanol added to the RM and MM medium is 8% and 3% (v/v), respectively.

E. coli DH5α was used for plasmid construction in this study. All *E. coli* strains were cultured in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, and 1.5% agar for solid) at 37°C, 250 rpm. When required, 100 and 200 μg/ml spectinomycin was added for *E. coli* or *Z. mobilis*, respectively.

Genetic manipulation and recombinant strain construction

The shuttle plasmid pEZ15A was used for *hfq* overexpression in *Z. mobilis* ZM4 with its native promoter and the recombinant strain named ZM4-*hfq*. The plasmid pL2R was used for *hfq* deletion using the native Type I-F CRISPR-Cas system of *Z. mobilis* (Zheng et al., 2019). Plasmid pEZ15A was transformed into the wild type and the *hfq* deletion strain, which named ZM4 and ZM4-Δ*hfq*, respectively.

For the gene-deficient mutant construction, interference plasmids were initially constructed with spacer. The spacer was designed to bear the entire 32-bp sequences containing a 5′-CCC-3′ PAM. The editing plasmid pL2R was digested with *Bsa* I at 37°C overnight. Oligonucleotides were annealed by first heating the reaction mixture to 95°C for 5 min and subsequently cooling down gradually to room temperature. The annealed spacer and the digested linear DNA vector were enzymatically linked with T4 ligase at 18°C overnight, and then transformed into DH5α competent cells to generate genome engineering plasmids.

For all plasmid construction, primers were designed to contain a region of 15–25 nucleotides that overlap with adjacent DNA fragments and synthesized by TsingKe (Beijing, China) (Supplementary Table S2). All plasmids were assembled by Gibson assembly method (Gibson et al., 2009). Gene and vector fragments amplified by primer pairs were purified and then ligated through the T5 exonuclease (NEB, WA, United States). The gene fragment was mixed with the vector at a 3:1 M ratio, 0.5 U T5 exonuclease (NEB, MA, United States) and 0.5 μl buffer 4 (NEB, MA, United States) were then added, with ddH_2O added to the final volume of 5 μl. All reagents were mixed and incubated on ice for 5 min before being added to chemically competent *E. coli* cells. After ice incubation for 30 min and heat-shock for 45 s at 42°C, held it on ice for 2 min before 100 μl NZY medium (5.0 g/L yeast extract, 5 g/L NaCl, 1.2 g/L $MgCl_2$, 1.5 g/L $MgSO_4$, 3.6 g/L glucose, 10 g/L casein enzymatic hydrolysate NZ amine®) were added to the mixture above, and recovered at 37°C, 250 rpm. The cells were plated on LB agar plates containing 100 μg/ml spectinomycin, the recombinant was screened by colony PCR and confirmed by Sanger Sequencing (TsingKe Biotechnology, Beijing, China).

After identification, the recombinant plasmid was transformed into *Z. mobilis* competent cells *via* electroporation (0.1-cm electrode gap, 1600 V, 200 Ω, 25 μF) using a Gene Pulser® (Bio-Rad, CA, United States) following the method developed for *Z. mobilis* (Okamoto and Nakamura, 1992). The correct colonies were selected by colony PCR. After three to five generations purification on RM agar plates with 200 μg/ml spectinomycin supplementation, the correct recombinant was stored at −80°C.

Cell growth and fermentation analysis

Seed culture of *Z. mobilis* was prepared by reviving the frozen glycerol stocks in 50 ml flasks containing 40 ml RM medium. After culturing overnight without shaking to the mid-exponential phase, the seed culture was harvested and inoculated in 50 ml shake flasks containing 40 ml RM or MM medium with an initial $OD_{600\ nm}$ value of 0.1. During the fermentation, cell growth in terms of the absorbance value (OD_{600}) was measured at 600 nm by a spectrophotometer (UV-1800, AOE, China) at different time points. Samples were centrifuged at 12,000 rpm for 2 min, collected supernatants were passed through 0.22-μm filters and stored at −80°C for subsequent HPLC analysis if needed. Glucose and ethanol in the culture supernatant were determined using a HPLC system LC-20AD (Shimadzu, Japan) with an Aminex HPX-87H column (300 × 7.8 mm, Bio-Rad, CA, United States) at 65°C. Elution was performed with 5 mM H_2SO_4 at 0.5 ml/min.

RNA-seq and statistical analyses

Cell culture samples in different conditions were collected at the mid-exponential phase. According to standard Illumina protocols, RNA-Seq was carried out by GENEWIZ (Suzhou, China) with a

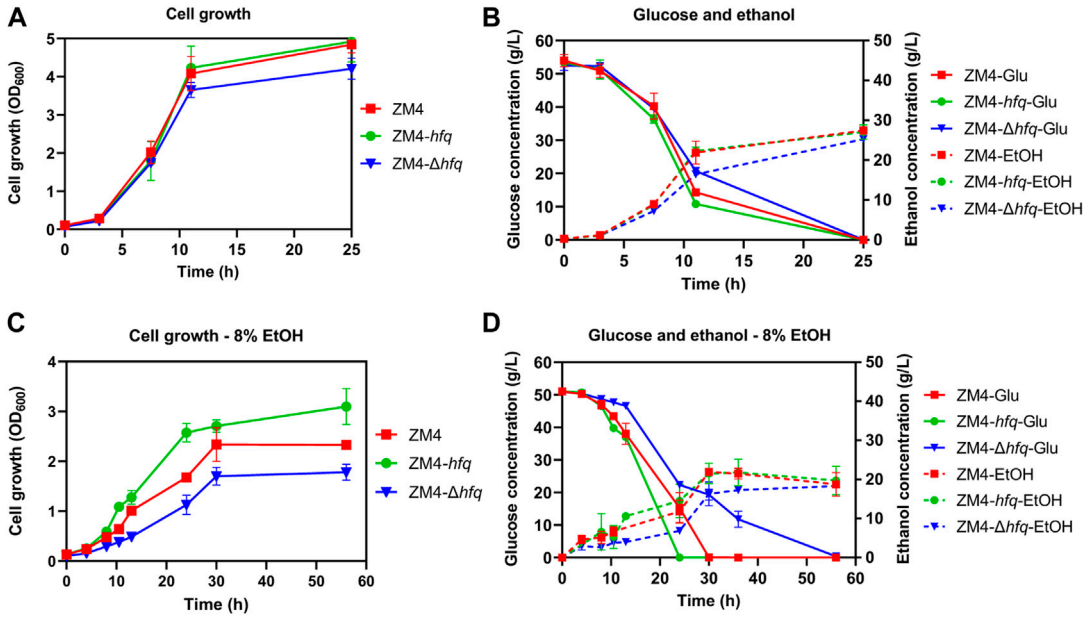


FIGURE 1
The cell growth, glucose consumption and ethanol production of *Z. mobilis* strains in the absence (**A,B**) or presence of 8% (v/v) ethanol (**C,D**). *Z. mobilis* wild-type (ZM4), *hfq* overexpression (ZM4-*hfq*). *hfq* deletion (ZM4-Δ*hfq*) strains were cultured in RM medium at 30°C. The error bars indicate standard deviations based on three replicates. Glu: glucose, EtOH: ethanol produced from glucose with the exogenous ethanol supplemented into the medium subtracted from the total ethanol amount in the medium.

library construction kit (NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina®) and an Illumina HiSeq 2000 instrument. After RNA-seq fastq data quality was evaluated by FastQC software (Babraham Bioinformatics, United Kingdom). Data were filtered and those passed the quality control were imported to the CLC Genomics Workbench (version 14.0) for reads trimming and RNA-Seq analysis to get the RPKM value (reads mapping to the genome per kilobase of transcript per million reads sequenced) of each gene. Genome sequence of *Z. mobilis* was used as the reference for RPKM calculation (Yang S. et al., 2018).

Gene expression normalization, analysis of variance (ANOVA), and hierarchical clustering analysis were conducted using JMP Genomics (Ver. 9.0, SAS Inc., NC, United States) to identify differentially expressed genes at different conditions. Significantly differentially expressed genes were identified with p-value ≤0.05 and $\log_2$-fold change ≥1 (significant induction) or ≤−1 (significant inhibition) as selection thresholds. The one-way ANOVA analysis was performed by JMP Genomics and only those with p-value ≤0.05 were considered for further analysis.

Measurement of intracellular ROS

Intracellular ROS levels were measured by 2′,7′-dichlorodihydro-fluorescein diacetate (H_2DCF-DA) (Beyotime Biotechnology, Hangzhou, China) (Yan et al., 2022). Cells were collected at the mid-log phase (total OD_{600} was 1.0) at 12,000 rpm for 1 min, washed with phosphate buffer saline (PBS) and then resuspended in 500 µl PBS with the final concentration of H_2DCF-DA at 100 µM. After incubating in darkness for 1 h at 30°C, 100 rpm, cells were collected, washed three times with PBS and resuspended in 300 µl PBS. The CytoFLEX FCM flow cytometry (Beckman Coulter, CA, United States) was used to detect the DCF fluorescence with excitation and emission wavelength of 488 nm and 525 nm, respectively. Cells with fluorescence intensities ranging 10^3–10^5 were selected and counted, and 20,000 cells were analyzed for each sample.

Yeast endoplasmic reticulum sequestration screening system

The plasmid pESE which is modified from pESD (Yi et al., 2013) was used and the TEV protease mediated Yeast Endoplasmic Reticulum Sequestration Screening (YESS) system was construction in *Z. mobilis* for the Protein-Protein interaction detection (Supplementary Figure S1). For the recombinant plasmid construction, two target genes and the vector fragment were PCR amplified and purified, and then assembled by the Golden gate assembly

TABLE 1 List of significantly differentially expressed genes in different functional categories between strain comparisons under 8% (v/v) ethanol treatment.

Locus ID	Gene	Product	ZM4-Δ*hfq* vs. ZM4	ZM4-*hfq* vs. ZM4	ZM4-Δ*hfq* vs. ZM4-*hfq*
Flagellar proteins					
ZMO0602	*motB1*	Flagellar motor protein Motb	—	—	—
ZMO0603	*motA1*	Flagellar motor protein Mota	—	—	—
ZMO0604	*flgL*	Flagellin domain protein	—	—	1.21
ZMO0605	*flgK*	Flagellar hook-associated protein FlgK	—	—	1.11
ZMO0606	-	Flagellar rod assembly protein FlgJ	1.34	—	1.79
ZMO0607	*flgI*	Flagellar P-ring protein	—	—	1.13
ZMO0608	*flgH*	Flagellar L-ring protein	—	—	1.08
ZMO0609	*flgG*	Flagellar basal-body rod protein FlgG	—	—	1.05
ZMO0613	*flgC*	Flagellar basal-body rod protein FlgC	—	—	1.08
ZMO0614	*flgB*	Flagellar basal-body rod protein FlgB	1.32	—	—
ZMO0629	*fliC*	Flagellin domain-containing protein	2.19	—	2.47
ZMO0631	-	Fis family sigma54-specific regulator	1.28	—	1.34
ZMO0632	*fliE*	Flagellar hook-basal body complex subunit FliE	1.49	—	2.10
ZMO0634	*fliF*	Flagellar M-ring protein FliF	1.24	—	2.06
ZMO0635	*fliG*	Flagellar motor switch protein FliG	1.46	—	2.06
ZMO0636	*fliH*	Negative regulator of FliI ATPase	—	−1.03	1.90
ZMO0637	*fliI*	Flagellum-specific ATP synthase FliI	—	—	1.34
Sulfur metabolism					
ZMO0003	*cysC*	Adenylyl-sulfate kinase	—	—	−2.21
ZMO0004	*cysN*	Sulfate adenylyltransferase large subunit	—	2.11	−2.70
ZMO0005	*cysD*	Sulfate adenylyltransferase small subunit	—	1.92	−2.68
ZMO0007	*cysH*	Phosphoadenosine phosphosulfate reductase	—	1.59	−1.73
ZMO0008	*cysI*	Sulfite reductase hemoprotein beta-component	—	1.64	−2.03
ZMO0009	*cysJ*	Sulfite reductase flavoprotein alpha chain	—	—	−1.89
ZMO1000	*metE*	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase	—	2.12	−2.56
ZMO1684	*serC*	Phosphoserine aminotransferase	—	1.09	—1.28
ZMO1685	*serA1*	D-3-phosphoglycerate dehydrogenase	—	1.55	—1.83
Heat shock response					
ZMO0016	*grpE*	GrpE protein	—	−1.10	1.06
ZMO0246	*hslV*	ATP-dependent protease subunit HslV	—	−2.06	2.21
ZMO0247	*hslU*	Heat shock protein atpase subunit HslU	—	−1.77	1.76
ZMO0376	*lon1*	Endopeptidase La	—	−1.21	1.62
ZMO0405	*clpA*	ATP-dependent Clp protease subunit ClpA	—	—	1.62
ZMO0660	*dnaK*	Chaperone protein DnaK	—	−1.51	1.60

(Continued on following page)

TABLE 1 (*Continued*) List of significantly differentially expressed genes in different functional categories between strain comparisons under 8% (v/v) ethanol treatment.

Locus ID	Gene	Product	ZM4-Δ*hfq* vs. ZM4	ZM4-*hfq* vs. ZM4	ZM4-Δ*hfq* vs. ZM4-*hfq*
ZMO0661	*dnaJ*	Chaperone protein DnaJ	—	−2.08	1.65
ZMO0948	*clpP*	ATP-dependent Clp protease subunit ClpP	—	—	1.27
ZMO0949	*clpX*	ATP-dependent Clp protease subunit ClpX	—	—	1.17
ZMO0989	*ibpA*	Heat-shock protein IbpA	—	−2.96	3.06
ZMO1424	*clpB*	ATP-dependent chaperone ClpB	—	−2.05	2.53
ZMO1704	*lon2*	Protease La (LON) substrate-binding domain	—	−1.65	2.22
ZMO1928	*groES*	Chaperonin Cpn10	—	−2.06	2.39
ZMO1929	*groEL*	Chaperonin GroEL	—	−2.20	2.28

a, gene expression represented the $\log_2$-fold changes; —for not differentially expressed; red or blue for up-regulated or down-regulated, respectively.

strategy. The reaction included 0.05 mol gene fragments, 20 ng pESE, 1 μl T4 ligase buffer (Thermo), 0.2 μl T4 ligase (Thermo), 0.3 μl *Bsa* I (NEB), 0.1 μl BSA (Takara), and with ddH_2O added to a final volume of 10 μl. All reagents were mixed and incubated on 37°C for 3 h before being added to chemically competent *E. coli* cells. After being incubated on ice for 30 min, heat-shocked for 45 s at 42°C, and held on ice for 2 min, 100 μl of NZY medium was added into above mixture and recovered for at least 1 h at 37°C with shaking at 250 rpm. The cells were plated on LB agar plates containing 50 μg/ml ampicillin, the recombinants were then screened by colony PCR and confirmed by Sanger Sequencing (TsingKe Biotechnology, Beijing, China).

The correct recombinant plasmids were transformed to *S. cerevisiae* EBY100 competent cells and then plated on SD-UT agar plates (2% galactose, 0.67% yeast nitrogen base w/o amino acids, 1% casamino acids, 1.5% agar) at 30°C. Single colony was selected and cultivated in 25 ml tube containing 5 ml SD-UT medium for 14 h with shaking at 200 rpm. The seed culture was then harvested and transferred into 25 ml tube containing 5 ml SD-DT medium (2% glucose, 0.67% yeast nitrogen base w/o amino acids, 1% casamino acids, 1.5% agar) with an initial $OD_{600\,nm}$ value of 0.8. After induction at 30°C for 8 h, 10^6 cells were collected by centrifugation, washing two times with 150 μl PBS buffer at 3,000 rpm, 4°C for 1 min, and finally resuspended in 20 μl PBS. For Fab display analysis, cells were incubated with 0.1 μl anti-HA-FITC and/or 0.1 μl anti-FLAG-iFlor647 (GenScript, Nanjing, China) at 4°C for 15 min in darkness. After washed three times with PBS (pH 7.4) supplemented with 0.5% BSA at 3,000 rpm, 4°C for 1 min, the sample was resuspended in 300 μl PBS (pH 7.4) with 0.5% BSA. Flow cytometry analysis was performed by the Beckman Coulter CytoFLEX (Brea, CA) equipped with 488 and 633 nm lasers and 525/40 and 660/20 nm band-pass filters.

Results and discussion

Hfq overexpression enhanced ethanol tolerance in *Z. mobilis*

To investigate the effect of *hfq* on ethanol tolerance in *Z. mobilis*, recombinant strains with *hfq* overexpression (ZM4-*hfq*) or *hfq* deletion (ZM4-Δ*hfq*) were constructed and evaluated in terms of cell growth, sugar utilization, and ethanol production. As shown in Figure 1A, the *hfq* overexpression strain ZM4-*hfq* grew normally without ethanol treatment and achieved a final OD_{600} value of 4.92 that was similar to the wild-type with OD_{600} value of 4.84. While, the *hfq* deletion strain ZM4-Δ*hfq* grew slightly slower than wild type with a final OD_{600} value of 4.20. Corresponding to the slower cell growth, ZM4-Δ*hfq* exhibited slower glucose consumption and ethanol production than the control and ZM4-*hfq* (Figure 1B). These results indicated that Hfq plays roles for normal cell growth even without ethanol stress inhibition.

Fermentation performance was further evaluated with 8% exogenous ethanol added into the media. Cell growth of ZM4 was effectively improved with *hfq* overexpression. ZM4-*hfq* grew into the stationary phase around 24 h and achieved a final OD_{600} value of 3.10, while wild-type strain had an OD_{600} value of 2.33 only after 56 h (Figure 1C). Consistently, *hfq* deletion in *Z. mobilis* significantly decreased the ethanol tolerance. The maximum biomass of strain ZM4-Δ*hfq* exposure to 8% ethanol was quite lower than that of the control with OD_{600} of 1.78 vs. 2.33. Corresponding to the cell growth, ZM4-*hfq* strain accelerated the glucose consumption and ethanol production, while strain ZM4-Δ*hfq* was in reverse. As shown in Figure 1D, all glucose was consumed by ZM4-*hfq* to achieve a maximum ethanol titer of 25.82 g/L within 24 h, almost 6 h faster than that of control, while 6.83 g/L residual glucose was still detected at 36 h by strain ZM4-Δ*hfq*. All these results indicated that *hfq* in

Z. mobilis contributes to the ethanol tolerance, just as its positive roles responding to other stresses such as furfural and acetic acid (Yang et al., 2010; Nouri et al., 2020).

Multiple genes regulated in the *hfq* recombinant strains under ethanol stress

Previous reports have shown that Hfq is associated with the regulation of numerous cellular pathways (Vogel and Luisi, 2011; Sharma et al., 2018; Han et al., 2019). To gain insights into the genes regulated by Hfq, and to illustrate the molecular mechanism of the positive role of Hfq overexpression on ethanol tolerance in *Z. mobilis*, RNA-Seq was employed to explore the global transcriptional differences of ZM4 and its *hfq* derivatives ZM4-*hfq* and ZM4-Δ*hfq* under different conditions of E0 (control, without exogenous ethanol treatment) and E8 (8% exogenous ethanol treatment) (Figure 1). Differentially expressed genes (DEGs) between different strains across different growth conditions were analyzed through analysis of variance (ANOVA) with a total of 628 DEGs identified (Supplementary Tables S3, S4).

The expression of *hfq* gene was examined first. As expected, the transcript of *hfq* was not detected in ZM4-Δ*hfq*, while the expression of *hfq* in ZM4-*hfq* had a 3.2-fold upregulation compared with the wild-type ZM4 (p-value ≤0.05) (Supplementary Tables S3, S4). The gene expression of *hfq* confirmed that recombinant strains ZM4-Δ*hfq* and ZM4-*hfq* were both constructed successfully, and the global gene expression was then further analyzed.

Consistent with previous studies (He et al., 2012; Yang et al., 2013), the presence of ethanol made many genes differentially expressed (E8 vs. E0) including 92 upregulated and 61 downregulated genes, which were mainly associated with cell wall/membrane biogenesis, metabolism, transcription, and general stress response. The effect of ethanol stress on three strains ZM4, ZM4-*hfq*, and ZM4-Δ*hfq* were further compared and analyzed. The results showed that exogenous ethanol stress regulated more than 200 genes in all three strains (Supplementary Table S3). This result suggested that despite of *hfq* overexpression, *Z. mobilis* remained challenged by the ethanol stress, which was consistent with the decreased cell growth and reduced biomass accumulation of ZM4-*hfq* in the presence of 8% ethanol compared with that without ethanol treatment (Figures 1A, C).

The detailed gene expression information of different strain comparisons was analyzed as well (Supplementary Table S4). Unlike the dramatic transcriptomic changes in response to ethanol stress, less differentially expressed genes were observed among strains. Under E0 condition without exogenous ethanol treatment, 58 genes were differentially expressed in ZM4-Δ*hfq* compared with ZM4, and 143 genes were differentially expressed when compared with ZM4-*hfq*. Meanwhile, only 30 genes were detected to be differentially regulated in ZM4-*hfq* compared with ZM4. When 8% exogenous ethanol was supplemented into RM (E8 condition), more differentially expressed genes were identified. Eighty genes were differentially expressed in ZM4-*hfq* compared with ZM4, while 77 genes and 181 genes were differentially expressed in ZM4-Δ*hfq* compared with ZM4 and ZM4-*hfq*, respectively. The result of more genes differentially expressed under E8 condition was consistent with the fermentation performance analyses that *hfq* recombinant strains were obviously different from the parental strain ZM4 under ethanol stress conditions (Figure 1). All these genes identified between strain comparisons could be closely associated with the positive role of Hfq on ethanol tolerance in *Z. mobilis*.

Hfq negatively regulated flagellar proteins in *Z. mobilis*

Bacterial flagellum is a complex and dynamic nanomachine appended on the cell body that provides motility (Soutourina and Bertin, 2003). Recent studies reported that these complex organelles also play an important role in bacterial survival, reproduction, and pathogenicity, such as adhesion to a variety of substrates, secretion of virulence factors, and formation of biofilms (Nedeljkovic et al., 2021). Gene expression without ethanol supplementation (E0) showed that 20 flagellar related genes (total 34 genes in *Z. mobilis*), including *flgBCEGHIJKL*, *flhA*, *fliCEFGHIKL*, and *motAB* encoding flagellar structure proteins, motor proteins and biosynthesis proteins were upregulated in ZM4-Δ*hfq* compared with the wild-type ZM4 or ZM4-*hfq*, and most of these genes were upregulated as well in the presence of 8% ethanol (E8) (Table 1, Supplementary Table S4).

Moreover, the expression of *fliA*, which encodes an alternative sigma factor specific for the flagellar regulons including *flgK* and *fliC* (Ohnishi et al., 1990), was two folds upregulated in ZM4-Δ*hfq* compared with the wild-type ZM4 with or without ethanol presence. These results collectively suggested that Hfq in *Z. mobilis* might constitutively mediate through the sigma factor FliA to negatively regulate the expression of flagellar proteins, which has been reported in other Gram-negative microbes, such as *Cronobacter sakazakii* (Kim et al., 2015). Previous transcriptomic analyses in *Z. mobilis* reported that flagellar related genes were downregulated to help conserve energy from cell motility for survival in the stressful conditions (He et al., 2012; Yang et al., 2013). Therefore, the less ethanol tolerance in ZM4-Δ*hfq* might be attributed to the limited energy resulting from the upregulation of the energy-costly flagellar assembly process without the negative regulation of Hfq.

In addition, gene expression in response to ethanol stress showed that above flagellar related genes were downregulated in

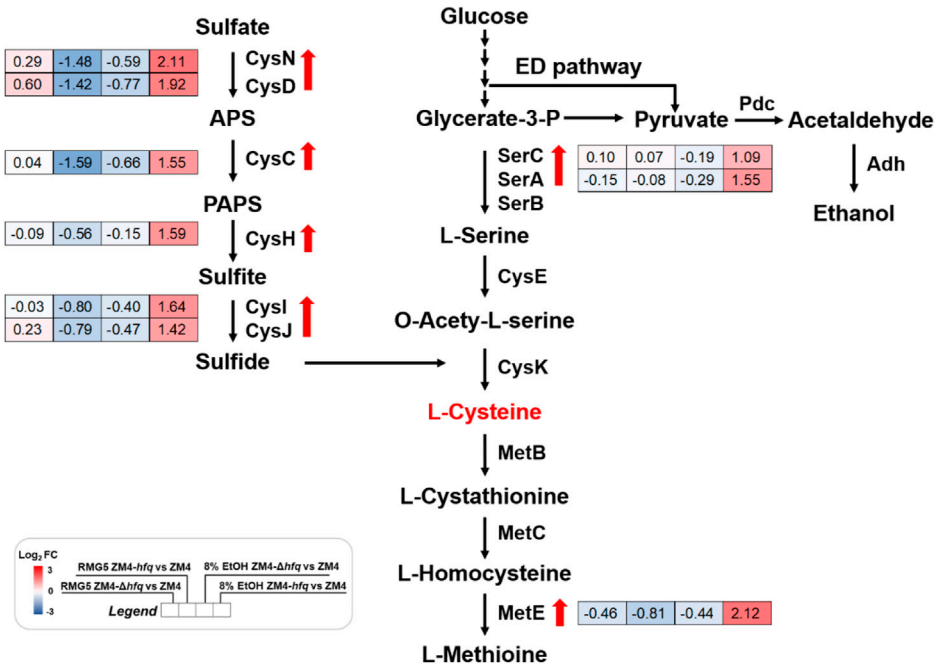


FIGURE 2
Transcriptional changes of genes involved in the sulfate assimilation and cysteine biosynthesis in *Z. mobilis* cultured in different conditions. Red and blue in the gene expression represent up-regulation and down-regulation respectively. Red arrows indicate the genes were significantly upregulated in ZM4-*hfq* compared with ZM4. APS: adenosine 5′-phosphosulfate, PAPS: 3′-phosphoadenylyl sulfate.

all three strains, especially in ZM4-Δ*hfq* with more than 2-fold repression (Supplementary Table S3). The downregulation of flagellar proteins in ZM4-Δ*hfq* indicated that other mechanisms of flagellar regulation may exist in *Z. mobilis* especially under the ethanol stress condition, and further study is needed to unravel the underlying mechanisms. Altogether, the gene expression data suggested that the negative regulation of Hfq on these flagellar related genes is beneficial but may not be sufficient for the ethanol tolerance in *Z. mobilis*.

Hfq mediated sulfur metabolism to eliminate the oxidative stress induced by ethanol stress

A gene cluster involved in sulfate assimilation and cysteine biosynthesis pathway was identified by comparing gene expression among strains under ethanol stress. Gene expression under ethanol treatment (E8) showed that all six genes associate with sulfate assimilation, *cysCND* (*ZMO0003-0005*) and *cysHIJ* (*ZMO0007-0009*), were upregulated in ZM4-*hfq* compared with ZM4-Δ*hfq* or ZM4. Similar upregulation of gene expression in ZM4-*hfq* was observed in other genes involved in cysteine biosynthesis, including *serA1* (*ZMO1685*) and *serC* (*ZMO1684*) (Table 1; Figure 2, Supplementary Table S4). However, such upregulated gene cluster was not induced under the condition without ethanol stress (E0). These results indicated that the sulfur metabolism including the cysteine biosynthesis was induced under the ethanol stress condition by *hfq* overexpression in *Z. mobilis*. Previous study confirmed that cysteine supplementation enhanced the tolerance in *Z. mobilis* to different inhibitors including ethanol (Yan et al., 2022). Therefore, *hfq* overexpression in ZM4-*hfq* might regulate the sulfate assimilation and cysteine biosynthesis for enhanced ethanol tolerance.

To evaluate the hypothesis that *hfq* overexpression enhanced sulfur metabolism and therefore played a protective role against the ethanol stress, 4 g/L Na_2SO_4 as the substrate for sulfate assimilation was supplemented into the medium and the cell growths of ZM4 and its *hfq* recombinant strains were detected (Figure 3). To avoid the effect of other sulfur-containing nutrients in RM medium, MM medium was used (Yan et al., 2022). Consistent with the result in rich medium (Figure 1), *hfq* overexpression strain ZM4-*hfq* exhibited an enhanced tolerance to 3% ethanol in MM media and achieved a final OD_{600} of 1.51, clearly higher than the wild-type with OD_{600} of 1.33. While, the *hfq* deletion strain ZM4-Δ*hfq* had the lowest cell growth with the final OD_{600} of 0.93 among three strains (Figure 3A). When

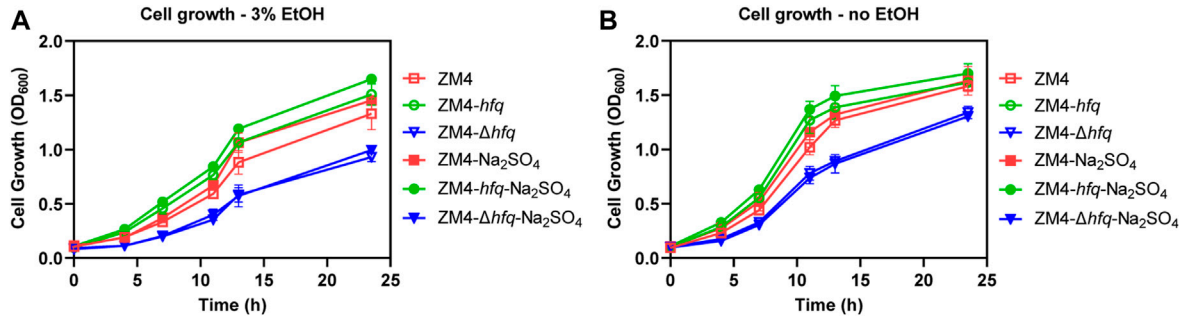


FIGURE 3
Effects of Na_2SO_4 addition on cell growth of *hfq* recombinant strains and wild type. 4 g/L Na_2SO_4 was added in the MM media and cell growths of *Z. mobilis hfq* recombinant strains were compared with the wild-type strain with **(A)** or without **(B)** 3% (v/v) ethanol treatment. EtOH represent ethanol.

Na_2SO_4 was added in the media, the cell growth of ZM4-*hfq* under ethanol treatment was further improved to a final OD_{600} of 1.65, while no influence was observed for *hfq* deletion strain ZM4-Δ*hfq*.

Moreover, an improvement responding to ethanol stress was detected as well in the wild-type ZM4 with the Na_2SO_4 supplementation, which suggested that the native gene expression of *hfq* in ZM4 would effectively regulate the sulfate assimilation process with Na_2SO_4 as the substrate. Since that MM medium is regarded as a stressful environment for *Z. mobilis* (Yang et al., 2014), Na_2SO_4 supplementation also slightly improved the cell growth of ZM4-*hfq* and ZM4 without ethanol treatment (Figure 3B). However, there is still no improvement observed for ZM4-Δ*hfq*. Taken together, Na_2SO_4 supplementation results suggested that the *hfq* overexpression in ZM4-*hfq* positively regulated the sulfate assimilation process, and Na_2SO_4 could be assimilated to deal with the stressful condition.

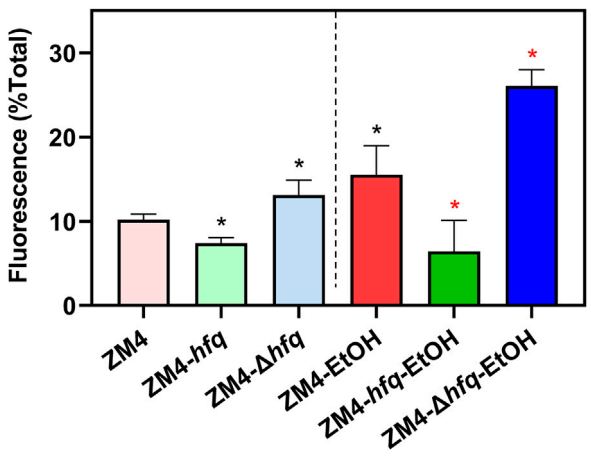


FIGURE 4
The effect of Hfq overexpression on reactive oxygen species (ROS) accumulation induced by ethanol in *Z. mobilis*. *Z. mobilis* strains were cultured in MM media with or without 3% (v/v) ethanol treatment, and EtOH represents the ethanol treatment. One-way ANOVA analysis was conducted for ROS detection with ZM4 (black asterisk, *p*-value <0.05) or ZM4-EtOH (red asterisk, *p*-value <0.05) condition as the control.

Previous studies in *E. coli*, *S. cerevisiae*, and *Z. mobilis* demonstrated that the sulfate assimilation was induced in response to stresses (Miller et al., 2009; Kanna and Matsumura, 2015; Zhang Y. et al., 2019). Recent study in *Z. mobilis* 8b further confirmed that cysteine supplementation in the growth media boosted glutathione synthesis or H_2S release effectively leading to the reduced accumulation of reactive oxygen species (ROS) induced by inhibitor stress (Yan et al., 2022). To evaluate the impact of upregulated sulfur metabolism on ROS for ethanol tolerance improvement in ZM4-*hfq*, the cellular ROS levels in *hfq* recombinant strains under ethanol treatment were further examined.

As shown in Figure 4, the intracellular ROS accumulation was detected in the wild-type ZM4 with ethanol treatment (10.19%–15.53%, *p*-value <0.05). Corresponding to the cell growth performance responding to ethanol stress (Figure 3), the intracellular ROS level was decreased by *hfq* overexpression in ZM4-*hfq* compared with the wild-type control ZM4 (15.53%–6.42%, *p*-value <0.05), while increased by *hfq* deletion in ZM4-Δ*hfq* (15.53%–26.07%, *p*-value <0.05). Similar alterations were detected in strain comparisons without ethanol presence (Figure 4). Collectively, the results suggested that *hfq* overexpression in *Z. mobilis* was involved in the positive regulation of sulfate assimilation and cysteine biosynthesis, and thus contributed to the effective elimination of the ROS induced by ethanol inhibition and finally enhanced the ethanol tolerance of *Z. mobilis*.

The major stress sigma factor RpoS of *E. coli* and *Salmonella* is the master regulator of oxidative stress response, and depends on Hfq for translation (Frohlich and

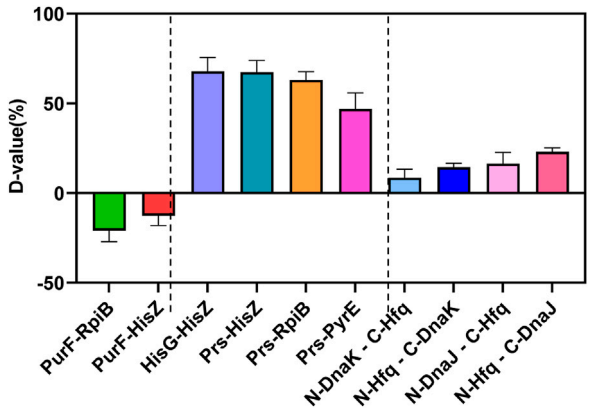


FIGURE 5
Protein-protein interactions of Hfq-DnaK and Hfq-DnaJ detected by the TEV protease mediated Yeast Endoplasmic Reticulum Sequestration Screening system. D-value: the difference of mono fluorescence signal between experimental samples and control pESE; PurF-RpiB, PurF-HisZ: negative control of protein-protein interaction; HisG-HisZ, Prs-HisZ, Prs-RpiB; Prs-PyrE: positive control of protein-protein interaction.

Gottesman, 2018). *Z. mobilis* does not possess a *rpoS* gene, but its genome contains other five sigma factors genes. RpoE (sigma E) in *Z. mobilis* potentially is a RpoS-like regulatory factor, and plays a critical role in various stress conditions including ethanol stress responses (Seo et al., 2005). Recent study in the facultative phototrophic bacterium *Rhodobacter sphaeroides* demonstrated that sigma factor RpoE is involved in a specific photo-oxidative stress response, and sRNA RSs0019 controlled by the sigma factor RpoE contributes to the balance of sulfate uptake and biosynthesis of sulfur-containing amino acids to guarantee appropriate cysteine and glutathione levels under the given environmental conditions (Hess et al., 2014). These evidences suggested a potential role of *rpoE* in the regulation of sulfur metabolism against the oxidative stress in *Z. mobilis*.

In this study, the gene expression of *rpoE* (*ZMO1404*) was not influenced in both *hfq* recombinant strains. However, Hfq may only influence the RpoE activity without affecting its protein levels just as observed in *R. sphaeroides* (Berghoff et al., 2011). In addition, recent study in *Z. mobilis* identified two sRNAs, Zms4 and Zms6, that might affect a wide range of sRNAs through modulating *hfq* expression indirectly and specifically co-regulate some pathways important to ethanol stress, including sulfate assimilation and cysteine biosynthetic processes (Han et al., 2020). All these results support the role of Hfq on sulfur metabolism in response to ethanol stress. However, considering the versatile function of Hfq (Dos Santos et al., 2019), further studies are necessary to define the regulation of Hfq in *Z. mobilis*.

Hfq involved in modulating RpoH-mediated cytoplasmic heat shock response

Z. mobilis can regulate general stress-response genes especially those in heat shock response to protect proteins from the damage caused by the stressful environments (Yang et al., 2014; Yi et al., 2015; Yang et al., 2020). Consistent with previous studies (He et al., 2012; Yang et al., 2013), many genes involved in the heat shock response were significantly upregulated in response to ethanol stress in ZM4, such as *grpE* (*ZMO0016*), *groES* (*ZM O 1928*), and *groEL* (*ZM O 1929*) (Supplementary Table S3). Further analysis discovered that more general stress response genes were induced by ethanol in ZM4-Δ*hfq*, including 11 genes involved in protein remodeling and reactivation, six genes involved in DNA replication and repair, and three genes responsive for oxidative stress. This was in accordance with the cell growth performance that ZM4-Δ*hfq* was more sensitive to ethanol stress (Figure 1). However, most of these stress response genes were significantly repressed by ethanol in strain ZM4-*hfq*.

Moreover, the gene expression in the strain comparisons showed that these stress response genes, especially genes encoding the heat shock protein (HSP) including *grpE* (*ZMO0016*), *dnaK* (*ZMO0660*), *dnaJ* (*ZMO0661*), *groES* (*ZMO1928*), and *groEL* (*ZMO1929*) were significantly downregulated in ZM4-*hfq* compared with the wild-type ZM4 or ZM4-Δ*hfq* under 8% ethanol (Table 1; Supplementary Table S4). These results collectively suggested that the stress responses induced by ethanol might be effectively alleviated by *hfq* overexpression in ZM4-*hfq*. Therefore, the strain did not have to upregulate its energy-costly protein repair system to conserve energy for ethanol stress responses and cell growth.

In many bacteria, RpoH (Sigma 32) is associated with the stress responses especially the heat shock proteins (HSPs) (Krajewski et al., 2014; Roncarati and Scarlato, 2017). Corresponding to the upregulation of HSPs, the expression of *rpoH* (*ZMO0749*) was induced two folds in ZM4-Δ*hfq* compared with ZM4 under ethanol treatment (Supplementary Table S4). Interestingly, recent study in *Z. mobilis* reported that *rpoH* overexpression led to a decrease in ethanol tolerance (Benoliel et al., 2020). Such evidence suggested that RpoH as well as HSPs at specific concentrations is required for *Z. mobilis* to respond to ethanol stress, and *rpoH* upregulation could make *Z. mobilis* sensitive to ethanol stress as observed in ZM4-Δ*hfq* (Figure 1).

Different from the upregulation of *rpoH* in ZM4-Δ*hfq*, the expression of *rpoH* was not affected in ZM4-*hfq* compared with ZM4 under ethanol treatment, even though the RpoH-transcribed heat shock regulons were downregulated (Supplementary Table S4). As a result, Hfq was speculated to be involved in the regulation of heat shock response *via* the alternative sigma factor RpoH. Previous studies demonstrated that RpoH activity is modulated *via* the DnaK/J-mediated

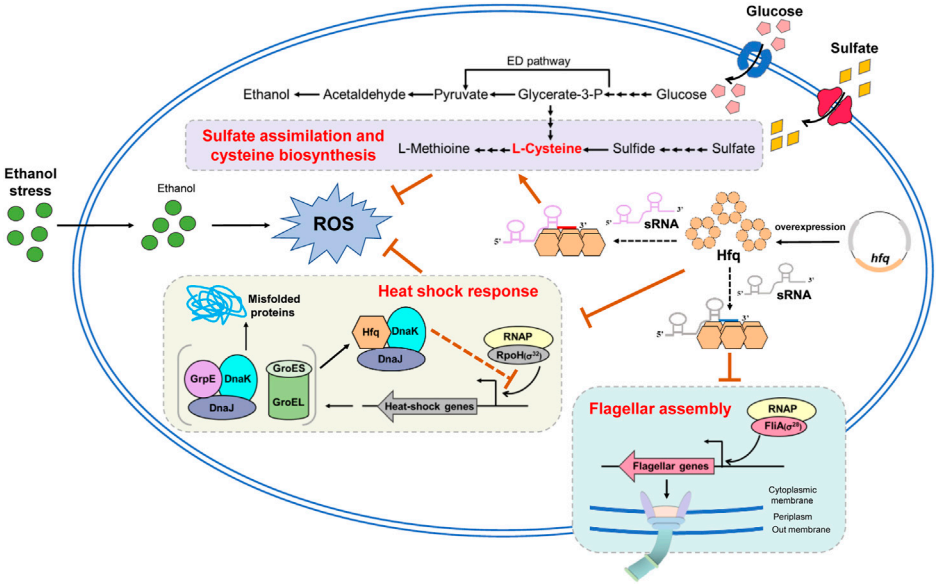


FIGURE 6
Proposed molecular mechanism of enhanced ethanol tolerance associated with *hfq* overexpression in *Z. mobilis*. Hfq overexpression helps negatively regulate the flagellar related genes and conserve energy for ethanol stress response and cell growth. Moreover, Hfq overexpression enhances sulfate assimilation and cysteine biosynthesis and eliminates the oxidative stress induced by ethanol inhibition. In addition, Hfq indirectly modulates the RpoH through interaction with DnaK/J and appropriately regulates the RpoH-mediated heat shock response under ethanol stress. ED pathway, Entner-Doudoroff pathway; ROS, reactive oxygen species. The orange arrows and orange bars represent upregulate and downregulate, respectively. Solid lines in this figure represent the confirmed regulations and dotted lines represent that are not yet confirmed.

negative feedback loop (Roncarati and Scarlato, 2017; Yura, 2019). In addition, Hfq was recently discovered to be capable to establish many protein-protein interactions in bacterial species (Yonekura et al., 2013; Dos Santos et al., 2019). Collectively, we speculated that *hfq* was associated with the indirect influence to RpoH activity *via* interaction with DnaK and/or DnaJ to help alleviate the cytoplasmic stress responses in *Z. mobilis* under ethanol stress.

Different methods have been developed to characterize the protein-protein interactions, such as co-immunoprecipitation (Co-IP), bacteria/yeast two-hybrid (B2H/Y2H), affinity-based techniques, and split protein complementation assays (Miura, 2018; Pichlerova and Hanes, 2021). Recently, Yi et al. (2013) developed a TEV protease mediated Yeast Endoplasmic Reticulum Sequestration Screening (YESS) system which can be easily applied in eukaryotic and prokaryotic species to explore the protein-protein interactions. In this system, the intracellular protein-protein interaction is converted to quantitative fluorescence signals through a split TEV protease-mediated proteolytic reaction, which possesses many advantages such as signal amplification, clean background and high efficiency. To illustrate the hypothesis above, the YESS system was established in *Z. mobilis* (Supplementary Figure S1) and then applied to detect the interaction between Hfq and DnaK/J.

As shown in Figure 5 and Supplementary Figure S2, the protein-protein interaction in YESS system was quantized by the D-value with flow cytometry. If the D-value >0, the protein pair has interaction, otherwise without interaction. Firstly, several known protein pairs of negative (PurF-RpiB, PurF-HisZ) and positive (HisG-HisZ, Prs-HisZ, Prs-RpiB, Prs-PyrE) ones were selected based on the UniProt and String database, and further evaluated. As expected, all the negative ones (PurF-RpiB, PurF-HisZ) displayed D-value <0, while the positive ones (HisG-HisZ, Prs-HisZ, Prs-RpiB, Prs-PyrE) showed D-value >0 (Figure 5 and Supplementary Figure S2). The results confirmed the YESS system was successfully established *in Z. mobilis*.

The protein pairs of Hfq-DnaK and Hfq-DnaJ were then characterized, which were linked with N-TEV or C-TEV respectively to avoid the structure interference in the protein interaction. Our results showed that the Hfq protein had direct interactions with DnaK and DnaJ in *Z. mobilis*, even though the interaction was not very strong with D-value from 8.45 to 22.98. The data indicated that Hfq was associated with the modulation of RpoH by direct interaction with DnaK/J. Taken together, the regulation of the RpoH-mediated cytoplasmic stress responses by Hfq contributed to the ethanol tolerance in *Z. mobilis*.

Conclusion

The accumulation of ethanol produced during fermentation is a bottleneck for bioethanol production improvement in *Z. mobilis*. This study exhibited that *hfq* overexpression enhanced ethanol tolerance and increased cell biomass in *Z. mobilis*. Combining transcriptomic analysis, biochemical, and genetics studies, the results identified that the improved ethanol tolerance by *hfq* overexpression is probably due to energy saving by downregulating flagellar biosynthesis and heat shock stress response proteins as well as reducing the ROS induced by ethanol stress *via* upregulating the sulfate assimilation and cysteine biosynthesis (Figure 6). This study gave a promising start for the characterization of the detailed regulatory mechanisms of Hfq in the stress responses in *Z. mobilis*. In addition, the YESS system established in this study provided an alternative approach to investigate the protein-protein interactions in *Z. mobilis*.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ Supplementary Material.

Author contributions

SY conceived and designed the experiments with inputs from all coauthors. YT performed the experiments with the help of YiW, QY, YZ, YaW, YY, MM, and MH. YT, XW, and SY analyzed the data. YT, XW, and SY wrote the manuscript with all authors conducted the extensive manuscript review. All authors contributed to data analyses, read, revised, and approved the final manuscript.

Funding

This work was supported by the National Key Technology Research and Development Program of China (2022YFA0911800 and 2018YFA0900300), the National Science Foundation of China (22108064, 21978071 and U1932141). The Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang Province (2018R01014), the Innovation Base for Introducing Talents of Discipline of Hubei Province (2019BJH021). Funding was also supported by Key Laboratory of Development and Application of Rural Renewable Energy, Ministry of Agriculture and Rural Affairs, China, and State Key Laboratory of Biocatalysis and Enzyme Engineering.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.1098021/full#supplementary-material

References

Benoliel, T., Rubini, M. R., de Souza Baptistello, C., Janner, C. R., Vieira, V. R., Torres, F. A., et al. (2020). Physiological effects of overexpressed sigma factors on fermentative stress response of *Zymomonas mobilis*. *Braz. J. Microbiol.* 51, 65–75. doi:10.1007/s42770-019-00158-3

Berghoff, B. A., Glaeser, J., Sharma, C. M., Zobawa, M., Lottspeich, F., Vogel, J., et al. (2011). Contribution of Hfq to photooxidative stress resistance and global regulation in *Rhodobacter sphaeroides*. *Mol. Microbiol.* 80, 1479–1495. doi:10.1111/j.1365-2958.2011.07658.x

Carreon-Rodriguez, O. E., Gutierrez-Rios, R. M., Acosta, J. L., Martinez, A., and Cevallos, M. A. (2019). Phenotypic and genomic analysis of *Zymomonas mobilis* ZM4 mutants with enhanced ethanol tolerance. *Biotechnol. Rep.* 23, e00328. doi:10.1016/j.btre.2019.e00328

Cho, S. H., Haning, K., Shen, W., Blome, C., Li, R. X., Yang, S. H., et al. (2017). Identification and characterization of 5 ' untranslated regions (5 ' UTRs) in *Zymomonas mobilis* as regulatory biological parts. *Front. Microbiol.* 8, 2432. doi:10.3389/fmicb.2017.02432

Dos Santos, R. F., Arraiano, C. M., and Andrade, J. M. (2019). New molecular interactions broaden the functions of the RNA chaperone Hfq. *Curr. Genet.* 65, 1313–1319. doi:10.1007/s00294-019-00990-y

Frohlich, K. S., and Gottesman, S. (2018). Small regulatory RNAs in the enterobacterial response to envelope damage and oxidative stress. *Microbiol. Spectr.* 6, 2018. doi:10.1128/microbiolspec.RWR-0022-2018

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi:10.1038/nmeth.1318

Han, R., Haning, K., Gonzalez-Rivera, J. C., Yang, Y., Li, R., Cho, S. H., et al. (2020). Multiple small RNAs interact to co-regulate ethanol tolerance in *Zymomonas mobilis*. *Front. Bioeng. Biotechnol.* 8, 155. doi:10.3389/fbioe.2020.00155

Han, Y., Chen, D., Yan, Y., Gao, X., Liu, Z., Xue, Y., et al. (2019). Hfq globally binds and destabilizes sRNAs and mRNAs in *Yersinia pestis*. *mSystems* 4, e00245–e00219. doi:10.1128/mSystems.00245-19

He, M. X., Wu, B., Qin, H., Ruan, Z. Y., Tan, F. R., Wang, J. L., et al. (2014). *Zymomonas mobilis*: A novel platform for future biorefineries. *Biotechnol. Biofuels* 7, 101. doi:10.1186/1754-6834-7-101

He, M. X., Wu, B., Shui, Z. X., Hu, Q. C., Wang, W. G., Tan, F. R., et al. (2012). Transcriptome profiling of *Zymomonas mobilis* under ethanol stress. *Biotechnol. Biofuels* 5, 75. doi:10.1186/1754-6834-5-75

Hess, W. R., Berghoff, B. A., Wilde, A., Steglich, C., and Klug, G. (2014). Riboregulators and the role of Hfq in photosynthetic bacteria. *RNA Biol.* 11, 413–426. doi:10.4161/rna.28035

Kanna, M., and Matsumura, Y. (2015). Effect of low-concentration furfural on sulfur amino acid biosynthesis in *Saccharomyces cerevisiae*. *J. Jpn. Pet. Inst.* 58, 165–168. doi:10.1627/jpi.58.165

Kim, S., Hwang, H., Kim, K. P., Yoon, H., Kang, D. H., and Ryu, S. (2015). *Hfq* plays important roles in virulence and stress adaptation in *Cronobacter sakazakii* ATCC 29544. *Infect. Immun.* 83, 2089–2098. doi:10.1128/IAI.03161-14

Krajewski, S. S., Joswig, M., Nagel, M., and Narberhaus, F. (2014). A tricistronic heat shock operon is important for stress tolerance of *Pseudomonas putida* and conserved in many environmental bacteria. *Environ. Microbiol.* 16, 1835–1853. doi:10.1111/1462-2920.12432

Miller, E. N., Jarboe, L. R., Turner, P. C., Pharkya, P., Yomano, L. P., York, S. W., et al. (2009). Furfural inhibits growth by limiting sulfur assimilation in ethanologenic *Escherichia coli* strain LY180. *Appl. Environ. Microbiol.* 75, 6132–6141. doi:10.1128/AEM.01187-09

Miura, K. (2018). An overview of current methods to confirm protein-protein interactions. *Protein Pept. Lett.* 25, 728–733. doi:10.2174/0929866525666180821122240

Nedeljkovic, M., Sastre, D. E., and Sundberg, E. J. (2021). Bacterial flagellar filament: A supramolecular multifunctional nanostructure. *Int. J. Mol. Sci.* 22, 7521. doi:10.3390/ijms22147521

Nouri, H., Moghimi, H., Marashi, S. A., and Elahi, E. (2020). Impact of *hfq* and *sigE* on the tolerance of *Zymomonas mobilis* ZM4 to furfural and acetic acid stresses. *PLoS One* 15, e0240330. doi:10.1371/journal.pone.0240330

Ohnishi, K., Kutsukake, K., Suzuki, H., and Iino, T. (1990). Gene *fliA* encodes an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*. *Molec. Gen. Genet.* 221, 139–147. doi:10.1007/BF00261713

Okamoto, T., and Nakamura, K. (1992). Simple and highly efficient transformation method for *Zymomonas mobilis*: Electroporation. *Biosci. Biotechnol. Biochem.* 56, 833. doi:10.1271/bbb.56.833

Pallach, M., Marchetti, R., Di Lorenzo, F., Fabozzi, A., Giraud, E., Gully, D., et al. (2018). *Zymomonas mobilis* exopolysaccharide structure and role in high ethanol tolerance. *Carbohydr. Polym.* 201, 293–299. doi:10.1016/j.carbpol.2018.08.072

Pichlerova, K., and Hanes, J. (2021). Technologies for the identification and validation of protein-protein interactions. *Gen. Physiol. Biophys.* 40, 495–522. doi:10.4149/gpb_2021035

Rogers, P. L., Jeon, Y. J., Lee, K. J., and Lawford, H. G. (2007). *Zymomonas mobilis* for fuel ethanol and higher value products. *Adv. Biochem. Eng. Biotechnol.* 108, 263–288. doi:10.1007/10_2007_060

Roncarati, D., and Scarlato, V. (2017). Regulation of heat-shock genes in bacteria: From signal sensing to gene expression output. *FEMS Microbiol. Rev.* 41, 549–574. doi:10.1093/femsre/fux015

Seo, J. S., Chong, H., Park, H. S., Yoon, K. O., Jung, C., Kim, J. J., et al. (2005). The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat. Biotechnol.* 23, 63–68. doi:10.1038/nbt1045

Sharma, I. M., Korman, A., and Woodson, S. A. (2018). The Hfq chaperone helps the ribosome mature. *EMBO J.* 37, e99616. doi:10.15252/embj.201899616

Soper, T. J., and Woodson, S. A. (2008). The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* 14, 1907–1917. doi:10.1261/rna.1110608

Soutourina, O. A., and Bertin, P. N. (2003). Regulation cascade of flagellar expression in Gram-negative bacteria. *FEMS Microbiol. Rev.* 27, 505–523. doi:10.1016/S0168-6445(03)00064-0

Tan, F. R., Wu, B., Dai, L. C., Qin, H., Shui, Z. X., Wang, J. L., et al. (2016). Using global transcription machinery engineering (gTME) to improve ethanol tolerance of *Zymomonas mobilis*. *Microb. Cell Fact.* 15, 4. doi:10.1186/s12934-015-0398-y

Thanonkeo, P., Laopaiboon, P., Sootsuwan, K., and Yamada, M. (2007). Magnesium ions improve growth and ethanol production of under heat or ethanol stress. *Biotechnology* 6, 112–119. doi:10.3923/biotech.2007.112.119

Updegrove, T. B., Zhang, A., and Storz, G. (2016). Hfq: The flexible RNA matchmaker. *Curr. Opin. Microbiol.* 30, 133–138. doi:10.1016/j.mib.2016.02.003

Vogel, J., and Luisi, B. F. (2011). Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9, 578–589. doi:10.1038/nrmicro2615

Wang, X., He, Q. N., Yang, Y. F., Wang, J. W., Haning, K., Hu, Y., et al. (2018). Advances and prospects in metabolic engineering of *Zymomonas mobilis*. *Metab. Eng.* 50, 57–73. doi:10.1016/j.ymben.2018.04.001

Yan, X., Wang, X., Yang, Y., Wang, Z., Zhang, H., Li, Y., et al. (2022). Cysteine supplementation enhanced inhibitor tolerance of *Zymomonas mobilis* for economic lignocellulosic bioethanol production. *Bioresour. Technol.* 349, 126878. doi:10.1016/j.biortech.2022.126878

Yang, S., Fei, Q., Zhang, Y., Contreras, L. M., Utturkar, S. M., Brown, S. D., et al. (2016). *Zymomonas mobilis* as a model system for production of biofuels and biochemicals. *Microb. Biotechnol.* 9, 699–717. doi:10.1111/1751-7915.12408

Yang, S. H., Franden, M. A., Wang, X., Chou, Y. C., Hu, Y., Brown, S. D., et al. (2020). Transcriptomic profiles of *Zymomonas mobilis* 8b to furfural acute and long-term stress in both glucose and xylose conditions. *Front. Microbiol.* 11, 13. doi:10.3389/fmicb.2020.00013

Yang, S., Pan, C., Hurst, Gregory B., Dice, Lezlee, Davison, Brian H., and Brown, S. D. (2014). Elucidation of *Zymomonas mobilis* physiology and stress responses by quantitative proteomics and transcriptomics. *Front. Microbiol.* 5, 246. doi:10.3389/fmicb.2014.00246

Yang, S., Pan, C., Tschaplinski, T. J., Hurst, G. B., Engle, N. L., Zhou, W., et al. (2013). Systems biology analysis of *Zymomonas mobilis* ZM4 ethanol stress responses. *PLoS One* 8, e68886. doi:10.1371/journal.pone.0068886

Yang, S., Pelletier, D. A., Lu, T. Y., and Brown, S. D. (2010). The *Zymomonas mobilis* regulator *hfq* contributes to tolerance against multiple lignocellulosic pretreatment inhibitors. *BMC Microbiol.* 10, 135. doi:10.1186/1471-2180-10-135

Yang, S., Vera, J. M., Grass, J., Savvakis, G., Moskvin, O. V., Yang, Y., et al. (2018a). Complete genome sequence and the expression pattern of plasmids of the model ethanologen *Zymomonas mobilis* ZM4 and its xylose-utilizing derivatives 8b and 2032. *Biotechnol. Biofuels* 11, 125. doi:10.1186/s13068-018-1116-x

Yang, Y., Hu, M., Tang, Y., Geng, B., Qiu, M., He, Q., et al. (2018b). Progress and perspective on lignocellulosic hydrolysate inhibitor tolerance improvement in *Zymomonas mobilis*. *Bioresour. Bioprocess.* 5, 6. doi:10.1186/s40643-018-0193-9

Yi, L., Gebhard, M. C., Li, Q., Taft, J. M., Georgiou, G., and Iverson, B. L. (2013). Engineering of TEV protease variants by yeast ER sequestration screening (YESS) of combinatorial libraries. *Proc. Natl. Acad. Sci. U. S. A.* 110, 7229–7234. doi:10.1073/pnas.1215994110

Yi, X., Gu, H., Gao, Q., Liu, Z. L., and Bao, J. (2015). Transcriptome analysis of *Zymomonas mobilis* ZM4 reveals mechanisms of tolerance and detoxification of phenolic aldehyde inhibitors from lignocellulose pretreatment. *Biotechnol. Biofuels* 8, 153. doi:10.1186/s13068-015-0333-9

Yonekura, K., Watanabe, M., Kageyama, Y., Hirata, K., Yamamoto, M., and Maki-Yonekura, S. (2013). Post-transcriptional regulator hfq binds catalase HPII: Crystal structure of the complex. *PLoS One* 8, e78216. doi:10.1371/journal.pone.0078216

Yura, T. (2019). Regulation of the heat shock response in *Escherichia coli*: History and perspectives. *Genes Genet. Syst.* 94, 103–108. doi:10.1266/ggs.19-00005

Zhang, K., Lu, X., Li, Y., Jiang, X., Liu, L., and Wang, H. (2019a). New technologies provide more metabolic engineering strategies for bioethanol production in *Zymomonas mobilis*. *Appl. Microbiol. Biotechnol.* 103, 2087–2099. doi:10.1007/s00253-019-09620-6

Zhang, Y., Ma, R., Zhao, Z., Zhou, Z., Lu, W., Zhang, W., et al. (2010). *irrE*, an exogenous gene from Deinococcus radiodurans, improves the growth of and ethanol production by a *Zymomonas mobilis* strain under ethanol and acid stress. *J. Microbiol. Biotechnol.* 20, 1156–1162. doi:10.4014/jmb.0912.12036

Zhang, Y., Vera, J. M., Xie, D., Serate, J., Pohlmann, E., Russell, J. D., et al. (2019b). Multiomic fermentation using chemically defined synthetic hydrolyzates revealed multiple effects of lignocellulose-derived inhibitors on cell physiology and xylose utilization in *Zymomonas mobilis*. *Front. Microbiol.* 10, 2596. doi:10.3389/fmicb.2019.02596

Zheng, Y., Han, J., Wang, B., Hu, X., Li, R., Shen, W., et al. (2019). Characterization and repurposing of the endogenous Type I-F CRISPR-Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res.* 47, 11461–11475. doi:10.1093/nar/gkz940

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Brief Research Report
PUBLISHED 23 December 2022
DOI 10.3389/fbioe.2022.1110513

OPEN ACCESS

EDITED BY
Zhenlin Han,
University of Hawaii at Manoa,
United States

REVIEWED BY
Riyi Xu,
Guangdong Academy of Sciences,
China
Quanfeng Liang,
Shandong University, China

*CORRESPONDENCE
Shihui Yang,
Shihui.Yang@hubu.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 28 November 2022
ACCEPTED 12 December 2022
PUBLISHED 23 December 2022

CITATION
Geng B, Liu S, Chen Y, Wu Y, Wang Y, Zhou X, Li H, Li M and Yang S (2022), A plasmid-free *Zymomonas mobilis* mutant strain reducing reactive oxygen species for efficient bioethanol production using industrial effluent of xylose mother liquor.
Front. Bioeng. Biotechnol. 10:1110513.
doi: 10.3389/fbioe.2022.1110513

A plasmid-free *Zymomonas mobilis* mutant strain reducing reactive oxygen species for efficient bioethanol production using industrial effluent of xylose mother liquor

Binan Geng[1], Shuyi Liu[1], Yunhao Chen[1], Yalun Wu[1], Yi Wang[1], Xuan Zhou[1], Han Li[1], Mian Li[2] and Shihui Yang[1]*

[1]State Key Laboratory of Biocatalysis and Enzyme Engineering, Environmental Microbial Technology Center of Hubei Province, School of Life Sciences, Hubei University, Wuhan, China, [2]Zhejiang Huakang Pharmaceutical Co., Ltd., Quzhou, Zhejiang, China

Genome minimization is an effective way for industrial chassis development. In this study, *Zymomonas mobilis* ZMNP, a plasmid-free mutant strain of *Z. mobilis* ZM4 with four native plasmids deleted, was constructed using native type I-F CRISPR-Cas system. Cell growth of ZMNP under different temperatures and industrial effluent of xylose mother liquor were examined to investigate the impact of native plasmid removal. Despite ZMNP grew similarly as ZM4 under different temperatures, ZMNP had better xylose mother liquor utilization than ZM4. In addition, genomic, transcriptomic, and proteomic analyses were applied to unravel the molecular changes between ZM4 and ZMNP. Whole-genome resequencing result indicated that an S267P mutation in the C-terminal of OxyR, a peroxide-sensing transcriptional regulator, probably alters the transcription initiation of antioxidant genes for stress responses. Transcriptomic and proteomic studies illustrated that the reason that ZMNP utilized the toxic xylose mother liquor better than ZM4 was probably due to the upregulation of genes in ZMNP involving in stress responses as well as cysteine biosynthesis to accelerate the intracellular ROS detoxification and nucleic acid damage repair. This was further confirmed by lower ROS levels in ZMNP compared to ZM4 in different media supplemented with furfural or ethanol. The upregulation of stress response genes due to the OxyR mutation to accelerate ROS detoxification and DNA/RNA repair not only illustrates the underlying mechanism of the robustness of ZMNP in the toxic xylose

Abbreviations: AhpC, Alkyl hydroperoxide reductase subunit C; ANOVA, analysis of variance; DEGs, differentially expressed genes; DEPs, differentially expressed proteins; ED, Entner-Doudoroff; Gor, glutathione reductase; Grx, glutaredoxin; GSH, glutathione; GSSG, glutathione (oxidized); LdH, actate dehydrogenase; NdH, NADH dehydrogenase; NGS, next-generation sequencing; OR, open reading frames; PHB, poly-3-hydroxybutyrate; ROS, reactive oxygen species; SNPs, single nucleotide polymorphisms; SOD, superoxide dismutase; TGS, third-generation sequencing; Ung, Uracil-DNA glycosylases; WGR, whole-genome resequencing.

mother liquor, but also provides an idea for the rational design of synthetic inhibitor-tolerant microorganisms for economic lignocellulosic biochemical production.

KEYWORDS

Zymomonas mobilis, plasmid-free, whole-genome resequencing (WGR), transcriptomics (RNA-Seq), proteomics, reactive oxygen species (ROS), oxyR, xylose mother liquor

Introduction

The production of sustainable, eco-friendly biofuels has become increasingly important in recent years due to the exhaustibility of fossil fuels and global climate change caused by excessive use of fossil fuels (Agarwal, 2007; Mills et al., 2009; Yang et al., 2018a). Bioethanol is an environmentally friendly and renewable liquid biofuel, and is one of the most promising alternatives to fossil fuels (Yang Y. et al., 2018). Lignocellulosic biomass is the most abundant, low-cost feedstock for bioethanol production (Balat et al., 2008; Zhang and Geng., 2012). However, these natural recalcitrant feedstocks are difficult to be directly utilized, therefore the processes of pretreatment and subsequent enzymatic hydrolysis are required to release sugars for microbial fermentation (Erdei et al., 2010; Yang Y., et al., 2018). During these processes, various toxic compounds such as acetic acid, furfural, HMF (5-hydroxymethyl-2-furaldehyde) are generated, which inhibit the substrate utilization, cell growth, and bioethanol production (Yang et al., 2018a). However, removal of these inhibitors isn't economically feasible due to the increased cost caused by additional processing steps and the potential loss of sugars for fermentation during this process (Parawira and Tekere, 2011; Jonsson et al., 2013). Therefore, it is crucial to develop and employ inhibitor-tolerant robust microorganisms against inhibitory compounds within the lignocellulosic hydrolysates for economic bioethanol production from lignocellulosic materials.

Microbial genome reduction and modification are important strategies for constructing promising cellular chassis (Leprince et al., 2012). The removal of many non-essential regions from a bacterial genome conferred beneficial traits such as genotypic stability and phenotypic validity (Li et al., 2016). For example, *E. coli* strain MDS42 with a 663 kb region of the genome deleted displayed two orders of magnitude higher electroporation efficiencies (Posfai et al., 2006). The mutant strain *E. coli* DGF-298 had a more stable genome, improved growth rate and cell density after deleting a 1,670 kb region of the genome. For the production of bioproducts, the genome-reduction strain *B. subtilis* BSK814G2 (814.4 kb deleted) accumulated 115.2 mg/L of guanosine which was 4.4-fold increase than the control strain. and *B. subtilis* BSK756T3 (756.8 kb deleted) accumulated 151.2 mg/L thymidine, showing a 5.2-fold increase compared to the control strain (Li et al., 2016). Additionally, after deleting 7.7% of the genome of *P. mendocina* NK-01, the medium-chain-length polyhydroxyalkanoates (PHAMCL) and alginate oligosaccharides (AO) yields of the resulting strain NKU421 increased by 114.8% and 27.8%, respectively (Fan et al., 2020). Therefore, genome reduction could be a viable approach to obtain inhibitor-tolerant strains for efficient utilization of lignocellulosic hydrolysates.

Zymomonas mobilis, a non-model ethanologenic Gram-negative bacterium, has many desirable industrial characteristics. For example, *Z. mobilis* has a highly specific rate of sugar uptake, exhibits very high ethanol tolerance (16% v/v) and high ethanol productivity (5.67 g/g/h), and can grow under a broad range of growth temperature (24–45°C) and pH (4.0–8.0) (Wang et al., 2018; Yang et al., 2021). *Z. mobilis* strains naturally use the anaerobic Entner-Doudoroff (ED) pathway efficiently to consume glucose for the production of ethanol and other biochemicals such as gluconic acid, levan, and sorbitol (Erzinger and Vitolo, 1996; Silbir et al., 2014). In addition, gene editing tools have been continuously developed in *Z. mobilis* including exogenous and native CRISPR-Cas genome editing tools (Shen et al., 2019; Zheng et al., 2019), and *Z. mobilis* has become a promising chassis for the economic production of lignocellulosic biofuels and biochemicals such as 2,3-butanediol, isobutanol, and poly-3-hydroxybutyrate (PHB) (Yang et al., 2016; Qiu et al., 2020; Li et al., 2022).

Since the first transcriptomic and metabolomic study on the effects of oxygen on *Z. mobilis* fermentation (Yang et al., 2009b), transcriptomics has been widely used to study the stress responses of *Z. mobilis* to ethanol, furfural, acetate, phenolic compounds, and other inhibitors (He et al., 2012b; Yang et al., 2014b). It is also used to explain the molecular mechanism of mutants obtained through mutagenesis or adaptative laboratory evolution such as acidic-pH-tolerant mutants (Yang Q et al., 2020), and reveal metabolic pathway changes, such as the cysteine biosynthesis pathway after cysteine supplementation (Yan et al., 2022). In addition, proteomics strategy has also been applied to provide an in-depth understanding of the stress responses of ethanol, acetate, and other inhibitors in *Z. mobilis* (Yang et al., 2013; Yang et al., 2014b; Chang et al., 2018). Multi-omics approaches including transcriptomic, proteomic, and whole-genome resequencing have also been widely

applied to investigate the changes and interrelationships of molecular components.

The genome of *Z. mobilis* model strain ZM4 includes a chromosome (2 Mb) and four native plasmids pZM32, pZM33, pZM36, and pZM39, which are 32–39 kb (Seo et al., 2005; Yang et al., 2009a; Yang et al., 2018b). These plasmids contain protein-coding genes involved in cellular defense, metabolism, and regulation (Yang et al., 2018b). For example, several genes involved in restriction and modification systems exist in pZM32. And genes encoding enzymes for metabolic functions are present in pZM33 including amidase, dehydrogenase, Acyl-CoA N-acyltransferase, NADPH-dependent oxidoreductase, and nucleoside triphosphate hydrolase. pZM36 contains several phage structure proteins. And genes encoding membrane-associated transporters, symporters, and porins are also identified in these plasmids, especially in pZM39. Addiction modules including toxin-antitoxin systems which have been reported to regulate stress adaptation and replicon persistence exist in all four plasmids (Unterholzner et al., 2013; Goeders & Van Melderen, 2014).

Although *Z. mobilis* is an excellent industrial microorganism for economic biochemical production from lignocellulosic materials, the capability of inhibitor tolerance can be further improved (He et al., 2012b; Yang et al., 2014a; Yi et al., 2015; Yang S et al., 2020). To examine whether the deletion of four native plasmids is beneficial for lignocellulosic ethanol production, we knocked out the cassette TetR-P*tet-cas12a* in ZMNP-Cas12a, a mutant with four native plasmids cured in previous work, to generate a plasmid-free mutant strain ZMNP in this study, and the ability of ZMNP to utilize lignocellulosic hydrolysate was further investigated.

Materials and methods

Strains, media, and culture conditions

Escherichia coli DH5α was cultured in Lysogeny Broth (LB, 10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract) at 37°C with shaking at 220 rpm. *Z. mobilis* ZM4 (ATCC 31821) was cultured in Rich Medium (RM, 50 g/L glucose, 10 g/L yeast extract, 2 g/L KH_2PO_4) and Minimal Medium (MM, 50 g/L glucose, 1 g/L KH_2PO_4, 1 g/L K_2HPO_4, 1 g/L $(NH_4)_2SO_4$, 0.5 g/L NaCl, 0.42 g/L $MgSO_4{\cdot}7H_2O$, 0.001 g/L calcium pantothenate) at 30°C with shaking at 100 rpm. When required, 50 μg/ml of chloramphenicol was added to the LB and RM medium, respectively. The final concentration of furfural or ethanol added to the RM medium was 2.60 g/L (RMF) and 47.36 g/L (RME), respectively. The final concentration of furfural or ethanol added to the MM medium was 1.25 g/L (MMF) and 20 g/L (MME), respectively.

Xylose mother liquor used in this study was provided by Zhejiang Huakang Pharmaceutical Co., Ltd. (Kaihua, Zhejiang, China). The original hydrolysate contained 216.87 g/L glucose, 103.98 g/L xylose, 4.44 g/L furfural, and 1.18 g/L acetic acid. The hydrolysate was diluted with ×10 RM^- medium (100 g/L yeast extract, 20 g/L KH_2PO_4) for ethanol fermentation.

Genetic manipulation and recombinant strain construction

Plasmid pL2R (Zheng et al., 2019) was used to knock out *cas12* gene in the chromosome. Meanwhile, gene *ZMO0038* was complemented at this position using the native Type I-F CRISPR-Cas system of *Z. mobilis* (Zheng et al., 2019). The spacer was 32 bp sequence which was immediately after a 5-NCC-3' PAM in *cas12a*. The oligonucleotides (TsingKe, Beijing, China) of spacer were 36 bp with 4 bp protruding sequences in 5'-end. The editing plasmid was constructed following the previous description(Zheng et al., 2019). Briefly, the targeting gRNA sequence was constructed by annealed two single-stranded oligonucleotides (Cas12a-gR-F: 5'-gaaatgcgttttgaactg attccgcagggtaaaacc, Cas12a-gR-R: 5'-gaacggtttttaccctgcggaatcagtt caaaacgca). Specifically, two single-stranded oligonucleotides were first heating at 95°C for 5 min and subsequently cooling down gradually to room temperature. Then the annealed spacer was ligated into *Bsa*I-linearized pL2R by T4 ligase at 22°C for 3 h. The resulting plasmid was named as pL2R-cas12a.

Gibson assembly method as described before (Li et al., 2022) was utilized for donor construction. Donor sequences including extra ~800 bp upstream sequence and downstream sequence of the candidate gene were amplified using Primer STAR polymerase (Takara, Japan) from the genomic DNA of *Z. mobilis* ZM4. The upstream sequence was amplified using the oligonucleotide primer pair 0038-US-F (5'-ggtcaccagctcaccgtctgt taggcgagaagggaaaggg) and 0038-US-R (5'-gttgggttgagccgcgatagt cgttaaatattcagatagacggagataataaacg), and the downstream sequence was amplified using the oligonucleotide primer pair 0038-DS-F (tcacgcccgacgccag) and 0038-DS-R (gctcgagatctgata tcactcaccctctggtgattgtcgat). The oligonucleotide primer pairs pL2R-FK-F (agtgatatcagatctcgagctcggtacccgg) and pL2R-FK-R (agacggtgagctggtgacct) were used to amplify the pL2R-cas12a vector. The upstream sequence and downstream sequence were then cloned into pL2R-cas12a vector by T5 exonuclease (NEB, WA, United States). The resulting plasmids were named as pL2R-cas12aD. The editing plasmid pL2R-cas12aD verified by colony PCR and Sanger sequencing were transformed into *Z. mobilis* ZMNP-Cas12a to construct the final strain ZMNP.

Electroporation transformation and recombinant strain selection

The editing plasmid pL2R-cas12aD was then transformed into *Z. mobilis* ZMNP-Cas12a competent cells (100 ng DNA with

50 μl competent cells) *via* electroporation using a Bio-Rad Gene Pulser (Bio-Rad, CA, United States). Immediately, the electroporated cells were transferred to 1 ml mating medium (50 g/L glucose, 10 g/L yeast extract, 5 g/L tryptone, 2.5 g/L $(NH_4)_2SO_4$, 0.2 g/L K_2HPO_4, 1 mM $MgSO_4$), and recovered at 30°C for 3 h. The cells were then spread on RM agar plates containing 50 μg/ml of chloramphenicol at 30°C for 2 days to isolate single colonies.

Flask fermentation and analytic methods

For the seed culture preparation, glycerol stock solution of *Z. mobilis* strains ZM4 and ZMNP were inoculated into 5-ml RM medium and then cultured at 30°C. The cultures were then transformed into 200 ml RM medium which was in 250-ml flasks, and then cultured at 30 C without shaking. The seed culture of mid-log phase was then inoculated into 50-ml shake flasks containing 40 ml RM, MM, RMF, MMF, RME, MME medium, 1/3 xylose mother liquor with an initial OD_{600} value of 0.1. During the different time points of fermentation, the OD_{600} value of cell culture was determined by an ultraviolet spectrophotometer UV 1800 (AOE, Shanghai, China). Simultaneously, samples collected at different time points were centrifuged at ×12,000 g for 2 min, and then the supernatants were filtered through a 0.22-μm filters and stored at −80°C for measuring the concentrations of glucose and ethanol by HPLC analysis.

For HPLC analysis, HPLC (LC-20AD, Shimadzu, Japan) with an Aminex HPX-87H column (Bio-Rad, CA, United States) were used to measure the glucose and ethanol concentrations at 65°C. At the same time, 5 mM H_2SO_4 was as mobile phase at a flow rate of 0.5 ml/min.

Whole-genome resequencing analysis

The sample ZMNP for whole-genome resequencing was collected and entrusted to GENEWIZ (Suzhou, China). The paired-end sequencing technology according to standard Illumina protocols by IgeneCode, Inc. (Beijing, China) was used in this work. The paired-end reads quality was checked using FastQC program (http://www.bioinformatics.babra.ham.ac.uk/projects/fastqc/). Data that passed the quality control were then mapped to the reference genome sequences of *Z. mobilis* ZM4 ATCC 31821 (GenBank accession No. of chromosome: NZ_CP023715, and plasmids: NZ_CP023716, NZ_CP023717, NZ_ CP023718, and NZ_CP023719) using the CLC Genomics Workbench (version 11.0) to identify the genomic variations. The objective mutations of the mutant strain ZMNP were obtained with the parental wild-type strain ZM4 as control. The mutation frequency which was more than 30% would be filtered.

RNA-Seq transcriptomic analysis

The method of transcriptomic study is the same as reported previously (Yang et al., 2018b; He et al., 2018). Briefly, cell cultured under RM and MM medium were collected at the mid-log phase. Then total RNA extraction was using TRIzol reagent (Invitrogen, CA, United States). Then rRNA was removed from total RNA by using Ribo-off rRNA Depletion Kit (Bacteria NR407). Subsequently, mRNA was interrupted to short fragments by adding the fragmentation buffer. After synthesizing the first strand cDNA using random hexamer-primers, the second-strand cDNA was synthesized using buffer, dNTPs, RNase H and DNA polymerase I, respectively. The sequencing library were constructed by connecting fragments and sequencing adapters. And the transcriptome data were sequenced based on the Illumina NovaSeq 6000 System.

RNA-Seq fastq data which passed the quality control by FastQC program were imported into CLC Genomics Workbench (version 14.0) for reads trimming and RNA-Seq analysis to get the RPKM values of each gene with *Z. mobilis* ZM4 ATCC 31821 (GenBank accession No. of chromosome: NZ_CP023715, and plasmids: NZ_CP023716, NZ_CP023717, NZ_CP023718, NZ_CP023719) and four native plasmids as the reference genome. JMP Genomics (version 9.0) was used to normalize gene expression, analyze variance (ANOVA) and hierarchical clustering to identify differentially expressed genes at different conditions. Differentially expressed genes were determined with a selection threshold of p-value ≤ 0.01 and $\log_2$-fold change≥ ± 1 (significant induction).

Protein sample preparation

The cells cultured in MM medium were collected and suspended into a 1 × RIPA buffer (×10: 150 mM NaCl, 10% NP40, 10% sodium deoxycholate, 10% SDS, 250 mM Tris–HCl, pH 7.6) at mid-log phase. Samples were grinded under low temperature for 5 min, and then ultrasonication on ice for 5 min. Subsequently, samples were kept at 4°C for 2 h and then centrifuged for 15 min at 4°C and ×12,000 g. 20 μl supernatant liquor containing extracted proteins of each sample was added into 96-well plate containing BCA buffer (Thermo Scientific, Rockford, United States). Then 96-well plate was shock at 37°C for 30 min, and the absorbance was detected at 562 nm (Thermo Scientific, CA, United States). The standard curve was then fitted, and the protein concentration of the corresponding sample is calculated. 100 μg of each sample was diluted to ~1 mg/ml with 1 × RIPA lysis buffer.

Then each sample was precipitated by acetone overnight at −20°C. Subsequently, the protein precipitation was collected by centrifugation for 10 min at 4°C and ×12,000 g and then rewashed with pre-chilled 80% acetone two times. The protein precipitation was re-dissolved, reduced, and alkylated. The

sequencing grade trypsin was then added at the ratio of 1:50 (wt: wt), and the digestion was run overnight at 37°C. The final protein precipitation was treated by sodium deoxycholate (SDC) cleanup and peptide desalting.

Nano-LC-MS/MS analysis

1 μg total peptides of each sample were separated and analyzed with a nano-UPLC (EASY- nLC1200) coupled to a Q Exactive HFX Orbitrap instrument (Thermo Fisher Scientific, CA, United States) with a nano-electrospray ion source. A reversed phase column (100 μm ID ×15 cm, Reprosil Pur 120 C18-AQ, 1.9 μm, Dr. Maisch) was used to separate different sizes of peptides. Mobile phases were H_2O with 0.1% FA, 2% ACN (phase A) and 80% ACN, 0.1% FA (phase B). Each ample was separated by executed with a 120 min gradient at 300 nL/min flow rate. Gradient B: 2%–5% for 2 min, 5%–22% for 88 min, 22%–45% for 26 min, 45%–95% for 2 min, 95% for 2 min. Data dependent acquisition (DDA) was performed in profile and positive mode with Orbitrap analyzer at a resolution of 120,000 (@200 m/z) and m/z range of 350–1,600 for MS1; For MS2, the resolution was set to 15,000 with a dynamic first mass. The automatic gain control (AGC) target for MS1 was set to 3E6 with max IT 50 ms, and 1E5 for MS2 with max IT 110 ms. The top 20 of most intense ions were fragmented by HCD with normalized collision energy (NCE) of 27%, and isolation window of 1.2 m/z. The dynamic exclusion time window was 45 s, single charged peaks and peaks with charge exceeding six were excluded from the DDA procedure.

Proteome Discoverer database search

Proteome Discoverer (PD) software (Version 2.4.0.305) and the built-in Sequest HT search engine were used to process Vendor's raw MS files. MS spectra lists were searched against their species-level UniProt FASTA databases (uniprot-Mus + musculus-10090-2020-10.fasta). Here, carbamidomethyl (C) as a fixed modification, oxidation (M) and acetyl (N-term) as variable modifications, and trypsin as proteases. A maximum of two missed cleavage(s) was allowed. The threshold of false discovery rate (FDR) was 0.01 for both PSM and peptide levels. Peptide identification was performed with an initial precursor mass deviation of up to 10 ppm and a fragment mass deviation of 0.02 Da. Unique peptide and Razor peptide were used for protein quantification and total peptide amount for normalization. All other parameters were reserved as default.

Measurement of intracellular ROS

Here, the measurement of intracellular ROS levels were using 2', 7'-dichlorodihydrofluorescein diacetate (H_2DCF-DA) (Beyotime Biotechnology, Hangzhou, China). When cells are cultured to mid-log phase, 0.6 OD_{600} cells were collected, and then washed with 1×phosphate-buffered saline (PBS) once. The pellets were re-suspended with 500 μl PBS, then added H_2DCF-DA at a final concentration of 100 μM. The mixture then incubated in darkness for 1 h at 30°C, 100 rpm. After that, cells were collected and washed three times with 1 × PBS. 300 μl 1 × PBS were used to re-suspended the pellets. The mixture then detected by the DCF fluorescence using CytoFLEX FCM flow cytometry (Beckman coulter, CA, United States). The excitation and emission wavelength were set up at 488 nm and 525 nm, respectively. Cells with fluorescence intensities ranging 10^3–10^5 were selected and counted, and at least 20,000 events were collected for each sample.

Results and discussion

Construction of *Z. mobilis* ZMNP and cell growth under different conditions

We previously constructed a mutant strain ZMNP-Cas12a that contains a cassette of TetR-P*tet-cas12a* that replaced *ZMO0038* locating between *ZMO0037* and *ZMO0039*. Compared to wild-type ZM4, ZMNP-Cas12a also lacks four native plasmids that were cured by CRISPR-Cas system. Here, we further constructed the mutant ZMNP with the cassette TetR-P*tet-cas12a* replaced with the native *ZMO0038* by native type I-F CRISPR system (Figure 1). The final mutant strain ZMNP only lacks four native plasmids compared to ZM4.

Temperature is one of the common physical factors affecting cell growth and microbial fermentation. Moreover, high-temperature ethanol fermentation has advantages of reduced pollution risk as well as cooling costs making it suitable for large-scale bioethanol fermentation (Li et al., 2021). Substrate is another factor that restricts the cost of microbial fermentation. Corncob is abundant with ca 250 million tons produced each year, which is an excellent cellulosic material for commercial xylose and xylitol production. However, the industrial effluent of xylose mother liquor accompanying the xylose and xylitol production becomes a problem. We evaluated cell growth and fermentation performance of ZM4 and ZMNP under different temperatures and different media including xylose mother liquor.

Cell growth and fermentation performance of ZMNP using glucose under different temperatures or using the xylose mother liquor were compared with ZM4, and the results demonstrated that ZMNP performed similarly to ZM4 using glucose with a growth rate of 0.525 h^{-1} *vs* 0.512 h^{-1} as well as ethanol productivity of 2.44 g/L h^{-1} *vs*

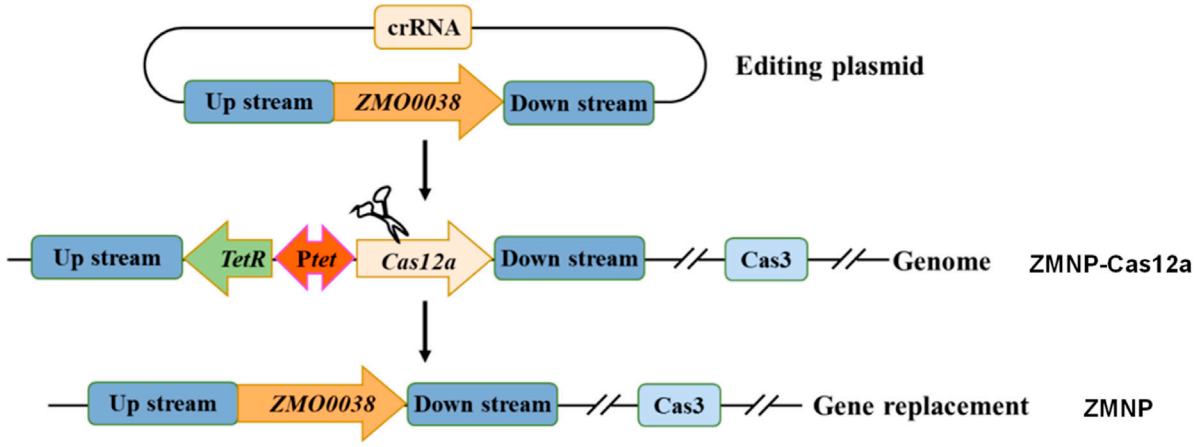


FIGURE 1
The scheme of ZMNP mutant construction by replacing cassette TetR-P*tet*-*cas12a* with *ZMO0038* in *Z. mobilis* ZMNP-Cas12a using the native type I-F CRISPR system.

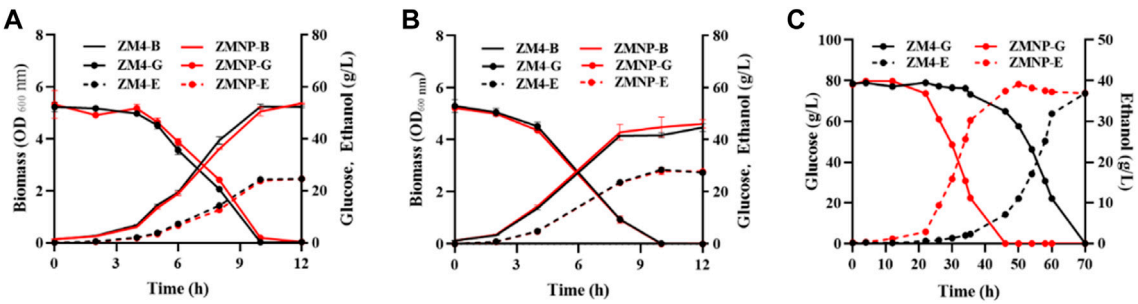


FIGURE 2
The effect of different temperatures and xylose mother liquor of *Z. mobilis* ZM4 and ZMNP. Growth, glucose consumption and ethanol production of ZMNP and ZM4 under 30°C **(A)** and 40°C **(B)**. Fermentation performance between ZMNP and ZM4 using xylose mother liquor **(C)**. Strains were cultured in 50-ml shake flasks containing 40 ml rich medium RM at 100 rpm. Three replicates were performed for the experiment. B: biomass, G: glucose, E: ethanol.

2.38 g/L h^{-1} when cultured under 30°C (Figure 2A), and with a growth rate of 0.395 h^{-1} *vs* 0.399 h^{-1} as well as ethanol productivity of 2.66 g/L h^{-1} *vs* 2.67 g/L h^{-1} when cultured under 40°C (Figure 2B). However, ZMNP did exhibit better glucose consumption and ethanol production than ZM4 when the toxic xylose mother liquor was used. The glucose utilization in ZMNP was 24 h faster than that observed for ZM4, and the maximum ethanol titers of 39.13 g/L and 36.91 g/L were achieved by ZMNP and ZM4, respectively (Figure 2C).

The normal growth and fermentation performance of ZMNP at high temperature and its excellent fermentation performance using xylose mother liquor exhibited that ZMNP could be an ideal chassis for biochemical production from lignocellulosic materials.

Genetic determinants of ZMNP for enhanced hydrolysate utilization

To determine the potential genetic determinants of ZMNP for increased utilization of xylose mother liquor, next-generation sequencing (NGS) and third-generation sequencing (TGS) technology were applied to identify the potential genetic changes in ZMNP. Combining the data collected from NGS and TGS, we assembled the genome of ZMNP to explore the potential genetic determinants for native plasmid deletion through comparative genomic analysis using the genome of parental strain ZM4 (ATCC 31821) as the reference (Yang et al., 2018b). The genome of ZMNP contains one circular chromosome of 2,058,754 bp only without plasmids that exist in the wild-type ZM4.

TABLE 1 Single-nucleotide polymorphisms (SNPs) in ZMNP compared to ZM4.

Locus	Gene	SNP	Frequency (%)	AA change	Product
225216	*ZMO0228*	C→T	100	Synonymous	putative polisoprenol-linked O-antigen translocase
481847	*ZMO0487*	C→T	99.64	Synonymous	HpcH/Hpal aldolase
641565	*ZMO0651*	G→A	99.74	D172N	Flagellar hook protein FliD
934794	*ZMO0915*	C→A	99.88	Synonymous	Copper-translocating P-type ATPase
1780057	*ZMO1733*	T→C	99.51	S267P	Transcriptional regulator OxyR

FIGURE 3
Overlay of the 3D structures of predicted OxyR mutant (Red) and wild-type OxyR protein (Green).

The whole-genome resequencing (WGR) results identified a total of five single nucleotide polymorphisms (SNPs) that occurred in open reading frames (ORF) in ZMNP (Table 1). Among these mutations, two missense mutations were identified in ZMNP located in the coding sequence region of *ZMO0651* (D172N) and *ZMO1733* (S267P), respectively. *ZMO0651* encodes a flagellar hook protein FliD, which is a flagellar cap at the distal end of the flagellar filament protecting the tip of the flagellum (Nedeljkovic et al., 2021). FliD also helps insert flagellin proteins repetitively to grow the flagellar filament (Cho et al., 2019). It was observed that the motility was defective in *fliD* mutants because the flagellin monomers were shed into the outside of cell causing the failure of flagellin polymerization and the form a functional flagellum, revealing that *fliD* was an essential factor in cell motility (Arora et al., 1998; Kim et al., 1999). We simulated the 3D structure of the FliD mutant, and it showed that this missense mutation didn't cause changes in the 3D structure (data not shown). Therefore, we speculate that this point mutation D172N mayn't affect its protein function.

The S267P mutation of OxyR is in LysR-substrate binding domain (IPR005119, 93-295 aa), which probably changes the binding affinity with its substrate (H_2O_2) due to the amino acid change from serine to proline, resulting in a protrusion at the 267 aa (Figure 3). Previous research reported that residue 266 aa was an important residue in reduced monomer conformation of OxyR (Anand et al., 2020), and beneficial to stabilize reduced tetramer *via* side-chain interactions (Anand et al., 2020). We speculated that the mutation at the 267 aa might affect the stabilization of the OxyR reduction state, making it more prone to transition to the oxidation state by altering the transcription initiation, and further respond to intracellular oxidative stress signals (Figure 1). This may be one of the reasons that ZMNP is better at utilizing xylose mother liquor.

Unravel the underlying mechanisms of hydrolysate tolerance of ZMNP through transcriptomic and proteomic studies

Overview of quantitative transcriptomics and proteomics

To illustrate the underlying genetic basis of the tolerance to xylose mother liquor in ZMNP, samples of ZMNP and wild-type strain ZM4 cultured at RM and MM were collected for RNA-Seq to explore the global transcriptional differences in ZMNP and ZM4. The differentially expressed genes (DEGs) were identified through analysis of variance (ANOVA) using strains and different media as variables. In addition to the genes in native plasmids, 1020 genes were identified by comparing both strains and media with p-value < 0.05 (Supplementary Table S1). Specifically, there were 55 and 87 DEGs comparing ZMNP with ZM4 at RM and MM conditions, respectively, reflecting the difference between strains under different media (Supplementary Table S2). 776 and 813 DEGs were also identified comparing MM with RM conditions of ZM4 and ZMNP, respectively, reflecting the effect of media for different strains (Supplementary Table S3). In addition to genes on native plasmids, 48 DEGs on genome were identified when took the complement of the DEGs under RM and MM media (Supplementary Table S2). Among these significantly DEGs, there were 29 genes upregulated and 19 genes downregulated in ZMNP compared with ZM4 (Supplementary Table S2). All these genes identified between strain comparisons were supposed to influence different functions in cell process, and then further analyzed.

We also used Quantitative proteomics to compare the differences in protein expression of *Z. mobilis* ZM4 and ZMNP in MM medium. We identified 14,801 peptides, which were matched to 1,425 unique proteins (Supplementary Table S4). Differentially expressed proteins (DEPs) which had more than 1.5-fold differences in abundance and less than 0.05 of p-value were determined and listed in Supplementary Material, showing the most abundant DEPs with 34 upregulated and 31 downregulated proteins of ZMNP (Supplementary Table S5).

Transcriptomic and proteomic profiling of *Z. mobilis* ZMNP compared to ZM4

Gene expression results showed that among the 19 downregulated genes in ZMNP, eight genes (*ZMO0383*, *ZMO0387*, *ZMO0388*, *ZMO0392*, *ZMO0397*, *ZMO0398*, *ZMO0930*, *ZMO1787*) encode hypothetical proteins, four genes encode levansucrase (*ZMO0374*, *ZMO0375*), levansucrase regulator (*ZMO0934*) and glycohydrolase secretase (*ZMO0064*), and two genes encode transporters (*ZMO0916*, *ZMO1457*) (Supplementary Table S2). Among them, the genes encoding levansucrase include *sacB* (*ZMO0374*) that catalyzes sucrose hydrolysis to glucose and fructo-oligosaccharides, and *sacC* (*ZMO0375*) that hydrolyzes sucrose to glucose and fructose (Gunasekaran et al., 1990). *zliE* (*ZMO0934*) encodes levansucrase regulator that stimulates the production of sucrose-hydrolyzing enzymes such as SacB and SacC (Kondo et al., 1994). *zliS* (*ZMO0932*) is a secretion-activating factor that contributes to the secretion of sucrose-hydrolyzing enzymes (Kondo et al., 1994). *oprB1* (*ZMO0064*) encodes a glucose porin that enables glucose entry into the periplasmic space and transports it to the cytoplasm *via* an ABC transport system (del Castillo et al., 2007). The downregulation of these genes indicate that the capacity of glycolysis, transport, and secretion of hydrolyzing enzymes was decreased in ZMNP, especially the hydrolysis of sucrose and the secretion of sucrose-hydrolyzing enzymes. The proteomic profiling results also showed that SacB and ZMO0916 were upregulated (Supplementary Table S5).

In addition, among the 29 upregulated genes in ZMNP, 15 genes were general stress response genes (Figure 4), especially the genes regulated by OxyR. It mainly includes 1) enzymes that are regulated by OxyR for synthesis of cellular antioxidants to remove peroxide; 2) enzymes that regulate and stabilize intracellular redox potential; 3) enzymes that repair DNA and RNA; and 4) enzymes associated with cysteine biosynthesis.

i) Upregulation of cellular antioxidants in ZMNP to remove peroxide for enhanced inhibitor tolerance. In the presence of ROS, the reduced form of OxyR could be transformed into oxidized OxyR by rapid formation of an intramolecular disulfide bond (Jo et al., 2015). In *E. coli*, the oxidized OxyR activates the expression of genes in response to oxidative stress (Ren et al., 2017) such as *sod*, *ahpC*, *yhjA*, *trx*, *grx*, *gor katG*, *ahpF*, and *hemF* (Charoenlap et al., 2005; Hishinuma et al., 2006; Zeller and Klug, 2006). The transition between reduced and oxidized OxyR also involves in the balance of GSH and GSSG, the multifunctional intracellular antioxidants with the reduced form and oxidized form, respectively. The oxidized OxyR is reduce by glutaredoxin (Grx) accompanied by the consumption of GSH and the production of GSSG. And GSSG could be transformed to reduced GSH by glutathione reductase (Gor), thus maintaining the balance between GSH and GSSG.

Our RNA-Seq results also demonstrated that six genes that are regulated by OxyR for oxidative stress response were upregulated. *ZMO1060*, encoding superoxide dismutase (SOD) to convert $O_2^{\bullet-}$ to O_2 and H_2O_2 (Zhao et al., 2021), was upregulated in ZMNP. This is similar to the results of previous studies that *sod* is positively regulated by OxyR (Zhou et al., 2021) and usually upregulated under stress conditions (Shatalin et al., 2011; Mironov et al., 2017). In addition, *ZMO1732* (*ahpC*), *ZMO1136* (*mauG*), and *ZMO1097* (*trx*) that are participated in the direct removal of H_2O_2 were significantly upregulated in ZMNP compared to ZM4. Alkyl

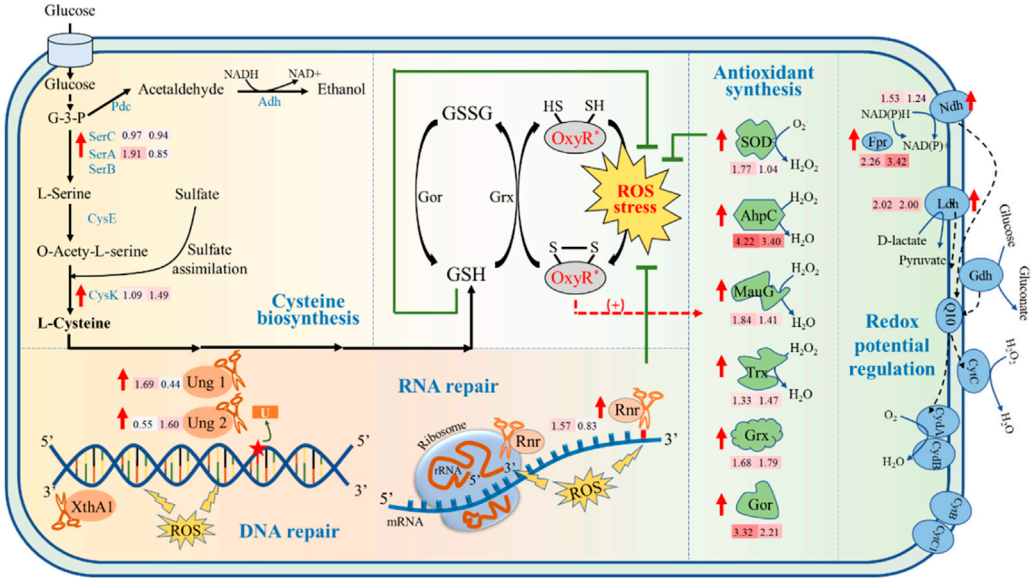


FIGURE 4
Potential molecular mechanism of inhibitor tolerance of ZMNP. AhpC: Alkyl hydroperoxide reductase, Fpr: Ferredoxin-NADP (+) reductase, Gor: Glutathione reductase, Grx: Glutaredoxin, GSH: Reduced glutathione, GSSG: Oxydized glutathione, G-3-P: Glyceraldehyde 3-Phosphate, Ldh: Lactate dehydrogenase, MauG: Cytochrome-c peroxidase, Ndh: FAD-dependent pyridine nucleotide-disulfide oxidoreductase, Rnr: Ribonuclease R, SOD: Superoxide dismutase, Trx: Thioredoxin, Ung: Uracil-DNA glycosylase, XthA1: Exodeoxyribonuclease III XthA.

hydroperoxide reductase subunit C (AhpC) proteins are the catalytic subunits of alkyl hydroperoxide reductases (Mongkolsuk et al., 2000), which are members of peroxidases to protect against H_2O_2. In ZMNP, the expression of *ZMO1732* was 10 folds and 18 folds higher than that in ZM4 when cultured under RM and MM, respectively. Protein MauG encoded by *ZMO1136* was 44% identity to YhjA of *E. coli* (NC_002695.2) by BLASTP analysis. It catalyzes the reduction of H_2O_2 to H_2O. The proteomic profiling results also showed that *ahpC* and *mauG* were upregulated. *ZMO1097* encoding thioredoxin (*trx*)-domain containing protein was also upregulated in ZMNP. Thioredoxin is a key antioxidant system that regulates protein dithiol/disulfide balance through its disulfide reductase activity to combat oxidative stress such as H_2O_2 (Lu and Holmgren, 2014). We speculate that *ZMO1097* was involved in the thioredoxin system and the overexpressed thioredoxin improved the Trx system's capacity to scavenge H_2O_2.

Our RNA-Seq results also demonstrated that the expression of *ZMO0753* (*grx*) encoding Grx that catalyzes the reaction of GSH to GSSG was significantly upregulated in both RM and MM in ZMNP. Furthermore, the expression of *ZMO1211* (*gor*) encoding Gor that catalyzes the reaction of GSSG to GSH was also upregulated in both RM and MM in ZMNP. The proteomic profiling results also showed that Grx and Gor were upregulated in ZMNP. The upregulation of both *ZMO0753* and *ZMO1211* indicates that the fast redox reaction in ZMNP contributes to the quick ROS removal. All six genes (*ZMO1060*, *ZMO1732*, *ZMO1136*, *ZMO1097*, *ZMO0753*, and *ZMO1211*) belong to OxyR regulon regulated by OxyR, revealing that the point mutation of OxyR in ZMNP might alter its three-dimensional structure and leads to the upregulation of genes involved in antioxidant regulation.

ii) Maintained intracellular redox potential in ZMNP for enhanced inhibitor tolerance. The microbial aerobic respiratory chain is an important source for cell metabolism to generate energy and to maintain the redox balance *in vivo*, which is related to cell metabolic processes such as ROS production and resistance to oxidative stress. Previous studies have shown that *Z. mobilis* possess a structural respiratory chain consisting mainly of type II NADH dehydrogenase (Ndh, *ZMO1113*), coenzyme Q10, cytochrome bd terminal oxidase (CydAB, *ZMO1571-ZMO1572*) (Kalnenieks et al., 2019), and D-lactate dehydrogenase (Ldh, *ZMO0256*) to contribute electrons to the respiratory chain. Our RNA-Seq data showed that *ZMO1113* and *ZMO0256* were upregulated in ZMNP, which accelerated the electron transport and reduced H_2O_2 and O_2 to H_2O. *ZMO1753* (Ferredoxin) that reduces $NADP^+$ to NADPH is also involved in the balance of the oxidation potential inside the cell. And the upregulation of *ZMO1753* in ZMNP could help maintain the intracellular redox potential.

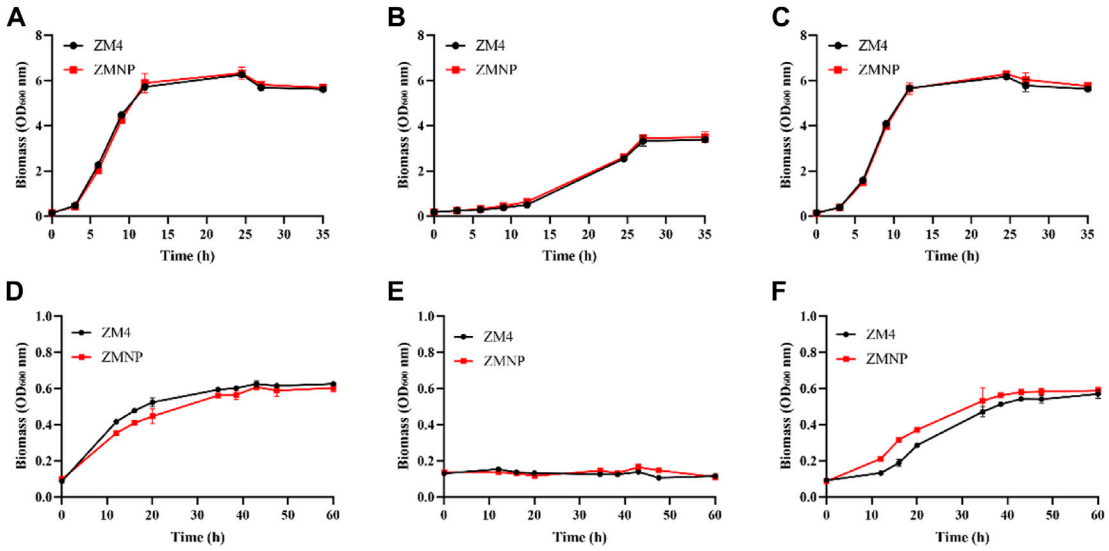


FIGURE 5
Cell growth of ZM4 and ZMNP responding to different inhibitors. RM **(A)** and RM supplemented with furfural (RMF) **(B)**, or ethanol (RME) **(C)**, as well as MM **(D)** and MM supplemented with furfural (MMF) **(E)**, or ethanol (MME) **(F)** were set up to evaluate the growth of ZM4 and ZMNP. RMF: 2.60 g/L furfural treatment in RM media; RME: 47.36 g/L ethanol treatment in RM media. MMF: 1.25 g/L furfural treatment in MM media, MME: 20 g/L ethanol treatment in MM media. At least two independent experiments were performed with similar results. Values are the mean of one representative experiment with three technical replicates. Error bars represent standard deviations.

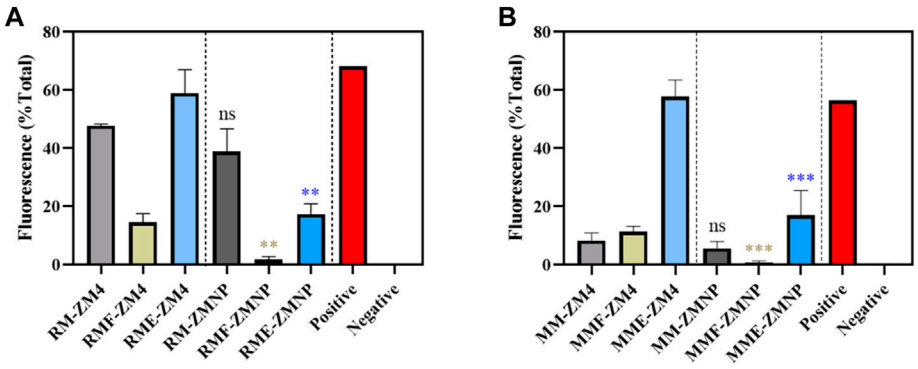


FIGURE 6
The ROS accumulation of ZM4 and ZMNP responding to different inhibitors. **(A)** ROS accumulation detection under RM with or without inhibitors. **(B)** ROS accumulation detection under MM with or without inhibitors. Data presented in the graphs are the mean ± SD of three replications. T-test analysis was conducted for ROS detection with RM/MM-ZM4 (black asterisk), RMF/MMF-ZM4 (green asterisk) or RME/MME (blue asterisk) condition as the control. ns represents no significant difference (p-value > 0.05), * represents a significant difference (0.01< p-value < 0.05), ** represents a significant difference (0.001 < p-value < 0.01), *** represents a significant difference (p-value < 0.001).

iii) Improved macromolecular repair in ZMNP for enhanced inhibitor tolerance. H_2O_2 produces hydroxyl radicals that oxidize the base and ribose parts of DNA, causing a variety of damages (Hutchinson, 1985; Imlay, 2013). Bacteria actively adopt a variety of oxidative stress measures to cope with oxidative stress, such as nucleotide excision, mismatch repair, and DNA double-strand breaks, which are initiated to protect genomic stability. The transcriptomic results showed that four genes associated with base excision repair were upregulated: *ZMO1114* (*ung1*), *ZMO1648* (*ung2*), *ZMO1401*(*xthA1*), and *ZMO1096* (*rnr*). Uracil-DNA glycosylases (Ung) are enzymes that cleave the bond between deoxyribose and mismatched uracil from

DNA (Krokan & Bjoras, 2013; Schormann et al., 2014). Two Ung enzyme genes, *ZMO1114* (*ung1*) and *ZMO1648* (*ung2*), were significantly upregulated in ZMNP, helping repair DNA damage caused by oxidative stress. Exonuclease III plays a key role in base excision repair which is a key repair mechanism to neutralize oxidative stress in DNA (Souza et al., 2006). In *Z. mobilis*, *ZMO1401* encodes a 3' to 5' exonuclease (Exonuclease III), and *ZMO1096* encodes a ribonuclease R to degrade RNA in the 3'-5' direction (Abula et al., 2022). Both *ZMO1401* and *ZMO1096* were significantly upregulated in ZMNP. Proteomic profiling results also showed that XthA1 was upregulated. These results suggested that ZMNP can protect DNA and RNA from damages caused by ROS through upregulation of these DNA/RNA repair proteins.

iv) Enhanced cysteine biosynthesis in ZMNP for inhibitor tolerance. Cysteine pool is crucial for microorganisms to defend against inhibitors. For example, cysteine is usually used for protein and GSH biosynthesis to protect cells against the oxidative stress (Hicks and Mullholland, 2018), and cysteine supplementation in the growth media helped reduce the toxicity of furfural and hydrolysates (Miller et al., 2009; Nieves et al., 2011; Yan et al., 2022). Gene expression results showed that three genes associated with cysteine synthesis were upregulated in ZMNP, including *ZMO1685* (*serA1*) and *ZMO1684* (*serC*) for L-serine synthesis, and *ZMO0748* (*cysK*) that could catalyze O-acety-L-serine and the sulfur assimilation product H_2S to synthesize L-cysteine. The upregulation of genes in cysteine biosynthesis pathway might increase the concentration of cysteine within cells to promotes GSH biosynthesis against ROS.

Upregulation of stress response genes in ZMNP helps reduce ROS level

Inhibitors such as furfural and high concentration of ethanol are well known to induce ROS accumulation. For example, furfural could damage various cellular components such as DNA, lipids, and proteins when ROS was accumulated in cells under stress conditions (Allen et al., 2010; Kim & Hahn, 2013). Ethanol is also one of the inhibitors of cell growth and metabolism in *Z. mobilis*. It can affect a wide range of cellular processes, such as DNA replication and recombination, DNA/RNA repair, transcriptional regulation, carbohydrate metabolism, cell wall/membrane biogenesis, terpenoid biosynthesis, respiratory chain, transport, and universal stress response (He et al., 2012a). All these inhibitors could contribute to an increase of intracellular ROS. *Z. mobilis* ZMNP upregulated global stress response genes that could help effectively reduce the intracellular ROS.

To examine whether the expression of genes related to ROS detoxification in ZMNP was upregulated, we determined the growth of ZMNP relative to ZM4 in both RM and MM. The concentrations of furfural that were supplemented into RM and MM were 2.60 g/L and 1.25 g/L, respectively, which were 47.36 g/L and 20 g/L for ethanol supplemented into RM and MM, respectively. Cell growth was similar when ZM4 and ZMNP were cultured in RM or RM media supplemented with furfural or ethanol (Figure 5).

Cells cultured in RM, RMF and RME for 3 h, and cultured in MM, MMF and MME for 12 h were collected to detect the ROS levels. The intracellular ROS levels were effectively decreased in ZMNP compared to ZM4 under all culture conditions (Figure 6). Compared with the control strain ZM4 (47.78%, 8.38%), ZMNP (38.87%, 5.70%) exhibited a slight decrease in intracellular ROS under RM and MM media, respectively. The difference of ROS levels between ZMNP and ZM4 were dramatic when cultured in inhibitor-supplemented media. ROS accumulation was detected when ZM4 and ZMNP were cultured in RMF (14.58% *vs* 1.80%) ($0.001 < p\text{-value} < 0.01$) and RME (58.97% *vs* 17.33%) ($0.001 < p\text{-value} < 0.01$), respectively. And the intracellular ROS levels were effectively decreased in ZMNP compared to ZM4 in MMF (11.44% *vs* 0.76%) ($0.001 < p\text{-value} < 0.01$) and MME (57.82% *vs* 16.98%) ($0.001 < p\text{-value} < 0.01$). Based on the result, the mechanism of efficient utilization of xylose mother liquor in ZMNP could be due to the overexpression of genes responsive for the general stress and intracellular ROS reduction.

Conclusion

Robust microorganisms are crucial for biochemical production from lignocellulosic materials. A plasmid-free ZMNP mutant of *Z. mobilis* was constructed in this study, which can efficiently utilize xylose mother liquor. Our multi-omics studies suggested that the S267P mutation of OxyR in ZMNP may help alter the expression of genes associated with global stress response under stress conditions, and ZMNP can reduce its intracellular ROS for efficient lignocellulosic bioethanol production. In addition, the molecular mechanism that OxyR regulates downstream genes to respond to the ROS of *Z. mobilis* proposed in this study can also guide the development of synthetic microbial cell factories for efficient lignocellulosic biochemical production.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

BG: Experimentations, Data analysis, and Manuscript writing. SL: Experimentations, Data analysis. YC: RNA-Seq Data analysis. YaW: Whole-genome resequencing Data analysis. YiW: Experimentations and Data analysis. XZ: Experimentations and Data analysis. HL: Data analysis. ML: Experimental materials, Data analysis, and manuscript revision. SY: Conceptualization, Funding acquisition, Supervision, Data analysis, Manuscript writing and submission.

Funding

This work was supported by the National Key Technology Research and Development Program of China (2022YFA0911800 and 2018YFA0900300), the National Natural Science Foundation of China (21978071 and U1932141). The Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang Province (2018R01014), the Innovation Base for Introducing Talents of Discipline of Hubei Province (2019BJH021). Funding was also supported by State Key Laboratory of Biocatalysis and Enzyme Engineering.

Conflict of interest

Author ML was employed by Zhejiang Huakang Pharmaceutical Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.1110513/full#supplementary-material

References

Abula, A., Yang, T., Zhang, Y., Li, T., and Ji, X. (2022). Enhancement of *Escherichia coli* ribonuclease R cytosine-sensitive activity by single amino acid substitution. *Mol. Biotechnol* [Epub ahead of print]. doi:10.1007/s12033-022-00533-w

Agarwal, A. K. (2007). Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Prog. Energy Combust. Sci.* 33 (3), 233–271. doi:10.1016/j.pecs.2006.08.003

Allen, S. A., Clark, W., McCaffery, J. M., Cai, Z., Lanctot, A., Slininger, P. J., et al. (2010). Furfural induces reactive oxygen species accumulation and cellular damage in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 3, 2. doi:10.1186/1754-6834-3-2

Anand, A., Chen, K., Catoiu, E., Sastry, A. V., Olson, C. A., Sandberg, T. E., et al. (2020). OxyR is a convergent target for mutations acquired during adaptation to oxidative stress-prone metabolic states. *Mol. Biol. Evol.* 37 (3), 660–667. doi:10.1093/molbev/msz251

Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S., and Ramphal, R. (1998). The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion. *Infect. Immun.* 66 (3), 1000–1007. doi:10.1128/iai.66.3.1000-1007.1998

Balat, M., Balat, H., and Öz, C. (2008). Progress in bioethanol processing. *Prog. Energy Combust. Sci.* 34 (5), 551–573. doi:10.1016/j.pecs.2007.11.001

Chang, D., Yu, Z., Ul Islam, Z., French, W. T., Zhang, Y., and Zhang, H. (2018). Proteomic and metabolomic analysis of the cellular biomarkers related to inhibitors tolerance in *Zymomonas mobilis* ZM4. *Biotechnol. Biofuels* 11, 283. doi:10.1186/s13068-018-1287-5

Charoenlap, N., Eiamphungporn, W., Chauvatcharin, N., Utamapongchai, S., Vattanaviboon, P., and Mongkolsuk, S. (2005). OxyR mediated compensatory expression between *ahpC* and *katA* and the significance of *ahpC* in protection from hydrogen peroxide in *Xanthomonas campestris*. *FEMS Microbiol. Lett.* 249 (1), 73–78. doi:10.1016/j.femsle.2005.06.002

Cho, S. Y., Song, W. S., Oh, H. B., Kim, H. U., Jung, H. S., and Yoon, S. I. (2019). Structural analysis of the flagellar capping protein FliD from *Helicobacter pylori*. *Biochem. Biophys. Res. Commun.* 514 (1), 98–104. doi:10.1016/j.bbrc.2019.04.065

del Castillo, T., Ramos, J. L., Rodriguez-Herva, J. J., Fuhrer, T., Sauer, U., and Duque, E. (2007). Convergent peripheral pathways catalyze initial glucose catabolism in Pseudomonas putida: Genomic and flux analysis. *J. Bacteriol.* 189 (14), 5142–5152. doi:10.1128/JB.00203-07

Erdei, B., Barta, Z., Sipos, B., Reczey, K., Galbe, M., and Zacchi, G. (2010). Ethanol production from mixtures of wheat straw and wheat meal. *Biotechnol. Biofuels* 3, 16. doi:10.1186/1754-6834-3-16

Erzinger, G. S., and Vitolo, M. (1996). *Zymomonas mobilis* as catalyst for the biotechnological production of sorbitol and gluconic acid. *Appl. Biochem. Biotechnol.* 131 (1-3), 787–794. doi:10.1385/abab:131:1:787

Fan, X., Zhang, Y., Zhao, F., Liu, Y., Zhao, Y., Wang, S., et al. (2020). Genome reduction enhances production of polyhydroxyalkanoate and alginate oligosaccharide in *Pseudomonas mendocina*. *Int. J. Biol. Macromol.* 163, 2023–2031. doi:10.1016/j.ijbiomac.2020.09.067

Goeders, N., and Van Melderen, L. (2014). Toxin-antitoxin systems as multilevel interaction systems. *Toxins (Basel)* 6 (1), 304–324. doi:10.3390/toxins6010304

Gunasekaran, P., Karunakaran, T., Cami, B., Mukundan, A. G., Preziosi, L., and Baratti, J. (1990). Cloning and sequencing of the *sacA* gene: Characterization of a sucrase from *Zymomonas mobilis*. *J. Bacteriol.* 172 (12), 6727–6735. doi:10.1128/jb.172.12.6727-6735.1990

He, M. X., Wu, B., Shui, Z. X., Hu, Q. C., Wang, W. G., Tan, F. R., et al. (2012a). Transcriptome profiling of *Zymomonas mobilis* under ethanol stress. *Biotechnol. Biofuels* 5 (1), 75. doi:10.1186/1754-6834-5-75

He, M. X., Wu, B., Shui, Z. X., Hu, Q. C., Wang, W. G., Tan, F. R., et al. (2012b). Transcriptome profiling of *Zymomonas mobilis* under furfural stress. *Appl. Microbiol. Biotechnol.* 95 (1), 189–199. doi:10.1007/s00253-012-4155-4

He, Q., Yang, Y., Yang, S., Donohoe, B. S., Van Wychen, S., Zhang, M., et al. (2018). Oleaginicity of the yeast strain *Saccharomyces cerevisiae* D5A. *Biotechnol. Biofuels* 11, 258. doi:10.1186/s13068-018-1256-z

Hicks, J. L., and Mullholland, C. V. (2018). Cysteine biosynthesis in Neisseria species. *Microbiology* 164 (12), 1471–1480. doi:10.1099/mic.0.000728

Hishinuma, S., Yuki, M., Fujimura, M., and Fukumori, F. (2006). OxyR regulated the expression of two major catalases, KatA and KatB, along with peroxiredoxin, AhpC in *Pseudomonas putida*. *Environ. Microbiol.* 8 (12), 2115–2124. doi:10.1111/j.1462-2920.2006.01088.x

Hutchinson, F. (1985). Chemical changes induced in DNA by ionizing radiation. *Prog. Nucleic Acid Res. Mol. Biol.* 32, 115–154. doi:10.1016/s0079-6603(08)60347-5

Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: Lessons from a model bacterium. *Nat. Rev. Microbiol.* 11 (7), 443–454. doi:10.1038/nrmicro3032

Jo, I., Chung, I. Y., Bae, H. W., Kim, J. S., Song, S., Cho, Y. H., et al. (2015). Structural details of the OxyR peroxide-sensing mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 112 (20), 6443–6448. doi:10.1073/pnas.1424495112

Jonsson, L. J., Alriksson, B., and Nilvebrant, N. O. (2013). Bioconversion of lignocellulose: Inhibitors and detoxification. *Biotechnol. Biofuels* 6 (1), 16. doi:10.1186/1754-6834-6-16

Kalnenieks, U., Balodite, E., and Rutkis, R. (2019). Metabolic engineering of bacterial respiration: High vs. low P/O and the case of *Zymomonas mobilis*. *Front. Bioeng. Biotechnol.* 7, 327. doi:10.3389/fbioe.2019.00327

Kim, D., and Hahn, J. S. (2013). Roles of the Yap1 transcription factor and antioxidants in *Saccharomyces cerevisiae*'s tolerance to furfural and 5-hydroxymethylfurfural, which function as thiol-reactive electrophiles generating oxidative stress. *Appl. Environ. Microbiol.* 79 (16), 5069–5077. doi:10.1128/aem.00643-13

Kim, J. S., Chang, J. H., Chung, S. I., and Yum, J. S. (1999). Molecular cloning and characterization of the *Helicobacter pylori fliD* gene, an essential factor in flagellar structure and motility. *J. Bacteriol.* 181 (22), 6969–6976. doi:10.1128/jb.181.22.6969-6976.1999

Kondo, Y., Toyoda, A., Fukushi, H., Yanase, H., Tonomura, K., Kawasaki, H., et al. (1994). Cloning and characterization of a pair of genes that stimulate the production and secretion of *Zymomonas mobilis* extracellular levansucrase and invertase. *Biosci. Biotechnol. Biochem.* 58 (3), 526–530. doi:10.1271/bbb.58.526

Krokan, H. E., and Bjoras, M. (2013). Base excision repair. *Cold Spring Harb. Perspect. Biol.* 5 (4), a012583. doi:10.1101/cshperspect.a012583

Leprince, A., van Passel, M. W., and dos Santos, V. A. (2012). Streamlining genomes: Toward the generation of simplified and stabilized microbial systems. *Curr. Opin. Biotechnol.* 23 (5), 651–658. doi:10.1016/j.copbio.2012.05.001

Li, R. X., Shen, W., Yang, Y. F., Du, J., Li, M., and Yang, S. H. (2021). Investigation of the impact of a broad range of temperatures on the physiological and transcriptional profiles of *Zymomonas mobilis* ZM4 for high-temperature-tolerant recombinant strain development. *Biotechnol. Biofuels* 14 (1), 146. doi:10.1186/s13068-021-02000-1

Li, Y., Wang, Y., Wang, R., Yan, X., Wang, J., Wang, X., et al. (2022). Metabolic engineering of *Zymomonas mobilis* for continuous co-production of bioethanol and poly-3-hydroxybutyrate (PHB). *Green Chem.* 24 (6), 2588–2601. doi:10.1039/d1gc04522a

Li, Y., Zhu, X., Zhang, X., Fu, J., Wang, Z., Chen, T., et al. (2016). Characterization of genome-reduced *Bacillus subtilis* strains and their application for the production of guanosine and thymidine. *Microb. Cell Fact.* 15, 94. doi:10.1186/s12934-016-0494-7

Lu, J., and Holmgren, A. (2014). The thioredoxin antioxidant system. *Free Radic. Biol. Med.* 66, 75–87. doi:10.1016/j.freeradbiomed.2013.07.036

Miller, E. N., Jarboe, L. R., Turner, P. C., Pharkya, P., Yomano, L. P., York, S. W., et al. (2009). Furfural inhibits growth by limiting sulfur assimilation in ethanologenic *Escherichia coli* strain LY180. *Appl. Environ. Microbiol.* 75 (19), 6132–6141. doi:10.1128/aem.01187-09

Mills, T. Y., Sandoval, N. R., and Gill, R. T. (2009). Cellulosic hydrolysate toxicity and tolerance mechanisms in *Escherichia coli*. *Biotechnol. Biofuels* 2, 26. doi:10.1186/1754-6834-2-26

Mironov, A., Seregina, T., Nagornykh, M., Luhachack, L. G., Korolkova, N., Lopes, L. E., et al. (2017). Mechanism of H_2S-mediated protection against oxidative stress in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 114 (23), 6022–6027. doi:10.1073/pnas.1703576114

Mongkolsuk, S., Whangsuk, W., Vattanaviboon, P., Loprasert, S., and Fuangthong, M. (2000). A xanthomonas alkyl hydroperoxide reductase subunit C (*ahpC*) mutant showed an altered peroxide stress response and complex regulation of the compensatory response of peroxide detoxification enzymes. *J. Bacteriol.* 182 (23), 6845–6849. doi:10.1128/jb.182.23.6845-6849.2000

Nedeljkovic, M., Sastre, D. E., and Sundberg, E. J. (2021). Bacterial flagellar filament: A supramolecular multifunctional nanostructure. *Int. J. Mol. Sci.* 22 (14), 7521. doi:10.3390/ijms22147521

Nieves, I. U., Geddes, C. C., Miller, E. N., Mullinnix, M. T., Hoffman, R. W., Fu, Z., et al. (2011). Effect of reduced sulfur compounds on the fermentation of phosphoric acid pretreated sugarcane bagasse by ethanologenic *Escherichia coli*. *Bioresour. Technol.* 102 (8), 5145–5152. doi:10.1016/j.biortech.2011.02.008

Parawira, W., and Tekere, M. (2011). Biotechnological strategies to overcome inhibitors in lignocellulose hydrolysates for ethanol production: Review. *Crit. Rev. Biotechnol.* 31 (1), 20–31. doi:10.3109/07388551003757816

Posfai, G., Plunkett, G., 3rd, Feher, T., Frisch, D., Keil, G. M., Umenhoffer, K., et al. (2006). Emergent properties of reduced-genome *Escherichia coli*. *Science* 312 (5776), 1044–1046. doi:10.1126/science.1126439

Qiu, M., Shen, W., Yan, X., He, Q., Cai, D., Chen, S., et al. (2020). Metabolic engineering of *Zymomonas mobilis* for anaerobic isobutanol production. *Biotechnol. Biofuels* 13, 15. doi:10.1186/s13068-020-1654-x

Ren, X., Zou, L., Zhang, X., Branco, V., Wang, J., Carvalho, C., et al. (2017). Redox signaling mediated by thioredoxin and glutathione systems in the central nervous system. *Antioxid. Redox Signal* 27 (13), 989–1010. doi:10.1089/ars.2016.6925

Schormann, N., Ricciardi, R., and Chattopadhyay, D. (2014). Uracil-DNA glycosylases-structural and functional perspectives on an essential family of DNA repair enzymes. *Protein Sci.* 23 (12), 1667–1685. doi:10.1002/pro.2554

Seo, J. S., Chong, H., Park, H. S., Yoon, K. O., Jung, C., Kim, J. J., et al. (2005). The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat. Biotechnol.* 23 (1), 63–68. doi:10.1038/nbt1045

Shatalin, K., Shatalina, E., Mironov, A., and Nudler, E. (2011). H_2S: A universal defense against antibiotics in bacteria. *Science* 334 (6058), 986–990. doi:10.1126/science.1209855

Shen, W., Zhang, J., Geng, B. N., Qiu, M. Y., Hu, M. M., Yang, Q., et al. (2019). Establishment and application of a CRISPR-Cas12a assisted genome-editing system in *Zymomonas mobilis*. *Microb. Cell Fact.* 18 (1), 162. doi:10.1186/s12934-019-1219-5

Silbir, S., Dagbagli, S., Yegin, S., Baysal, T., and Goksungur, Y. (2014). Levan production by *Zymomonas mobilis* in batch and continuous fermentation systems. *Carbohydr. Polym.* 99, 454–461. doi:10.1016/j.carbpol.2013.08.031

Souza, L. L., Eduardo, I. R., Padula, M., and Leitao, A. C. (2006). Endonuclease IV and exonuclease III are involved in the repair and mutagenesis of DNA lesions induced by UVB in *Escherichia coli*. *Mutagenesis* 21 (2), 125–130. doi:10.1093/mutage/gel006

Unterholzner, S. J., Poppenberger, B., and Rozhon, W. (2013). Toxin-antitoxin systems: Biology, identification, and application. *Mob. Genet. Elem.* 3 (5), e26219. doi:10.4161/mge.26219

Wang, X., He, Q. N., Yang, Y. F., Wang, J. W., Haning, K., Hu, Y., et al. (2018). Advances and prospects in metabolic engineering of *Zymomonas mobilis*. *Metab. Eng.* 50, 57–73. doi:10.1016/j.ymben.2018.04.001

Yan, X., Wang, X., Yang, Y., Wang, Z., Zhang, H., Li, Y., et al. (2022). Cysteine supplementation enhanced inhibitor tolerance of *Zymomonas mobilis* for economic lignocellulosic bioethanol production. *Bioresour. Technol.* 349, 126878. doi:10.1016/j.biortech.2022.126878

Yang, Q., Yang, Y., Tang, Y., Wang, X., Chen, Y., Shen, W., et al. (2020a). Development and characterization of acidic-pH-tolerant mutants of *Zymomonas mobilis* through adaptation and next-generation sequencing-based genome resequencing and RNA-Seq. *Biotechnol. Biofuels* 13, 144. doi:10.1186/s13068-020-01781-1

Yang, S., Franden, M. A., Brown, S. D., Chou, Y. C., Pienkos, P. T., and Zhang, M. (2014a). Insights into acetate toxicity in *Zymomonas mobilis* 8b using different substrates. *Biotechnol. Biofuels* 7 (1), 140. doi:10.1186/s13068-014-0140-8

Yang, S., Franden, M. A., Wang, X., Chou, Y. C., Hu, Y., Brown, S. D., et al. (2020b). Transcriptomic profiles of *Zymomonas mobilis* 8b to furfural acute and long-term stress in both glucose and xylose conditions. *Front. Microbiol.* 11, 13. doi:10.3389/fmicb.2020.00013

Yang, S., Franden, M. A., Yang, Q., Chou, Y. C., Zhang, M., and Pienkos, P. T. (2018a). Identification of inhibitors in lignocellulosic slurries and determination of their effect on hydrocarbon-producing microorganisms. *Front. Bioeng. Biotechnol.* 6, 23. doi:10.3389/fbioe.2018.00023

Yang, S. H., Mohagheghi, A., Franden, M. A., Chou, Y. C., Chen, X. W., Dowe, N., et al. (2016). Metabolic engineering of *Zymomonas mobilis* for 2, 3-butanediol production from lignocellulosic biomass sugars. *Biotechnol. Biofuels* 9 (1), 189. doi:10.1186/s13068-016-0606-y

Yang, S., Pan, C., Hurst, G. B., Dice, L., Davison, B. H., and Brown, S. D. (2014b). Elucidation of *Zymomonas mobilis* physiology and stress responses by quantitative proteomics and transcriptomics. *Front. Microbiol.* 5, 246. doi:10.3389/fmicb.2014.00246

Yang, S., Pan, C., Tschaplinski, T. J., Hurst, G. B., Engle, N. L., Zhou, W., et al. (2013). Systems biology analysis of *Zymomonas mobilis* ZM4 ethanol stress responses. *PLoS One* 8 (7), e68886. doi:10.1371/journal.pone.0068886

Yang, S., Pappas, K. M., Hauser, L. J., Land, M. L., Chen, G. L., Hurst, G. B., et al. (2009a). Improved genome annotation for *Zymomonas mobilis*. *Nat. Biotechnol.* 27 (10), 893–894. doi:10.1038/nbt1009-893

Yang, S., Tschaplinski, T. J., Engle, N. L., Carroll, S. L., Martin, S. L., Davison, B. H., et al. (2009b). Transcriptomic and metabolomic profiling of *Zymomonas mobilis* during aerobic and anaerobic fermentations. *BMC Genomics* 10, 34. doi:10.1186/1471-2164-10-34

Yang, S., Vera, J. M., Grass, J., Savvakis, G., Moskvin, O. V., Yang, Y., et al. (2018b). Complete genome sequence and the expression pattern of plasmids of the model ethanologen *Zymomonas mobilis* ZM4 and its xylose-utilizing derivatives 8b and 2032. *Biotechnol. Biofuels* 11, 125. doi:10.1186/s13068-018-1116-x

Yang, Y. F., Geng, B. N., Song, H. Y., He, Q. N., He, M. X., Bao, J., et al. (2021). Progress and perspectives on developing *Zymomonas mobilis* as a chassis cell. *Synthetic Biol. J.* 2 (1), 59–90. doi:10.12211/2096-8280.2020-071

Yang, Y., Hu, M., Tang, Y., Geng, B., Qiu, M., He, Q., et al. (2018c). Progress and perspective on lignocellulosic hydrolysate inhibitor tolerance improvement in *Zymomonas mobilis*. *Bioresour. Bioprocess.* 5 (1), 6. doi:10.1186/s40643-018-0193-9

Yi, X., Gu, H., Gao, Q., Liu, Z. L., and Bao, J. (2015). Transcriptome analysis of *Zymomonas mobilis* ZM4 reveals mechanisms of tolerance and detoxification of phenolic aldehyde inhibitors from lignocellulose pretreatment. *Biotechnol. Biofuels* 8, 153. doi:10.1186/s13068-015-0333-9

Zeller, T., and Klug, G. (2006). Thioredoxins in bacteria: Functions in oxidative stress response and regulation of thioredoxin genes. *Naturwissenschaften* 93 (6), 259–266. doi:10.1007/s00114-006-0106-1

Zhang, W., and Geng, A. (2012). Improved ethanol production by a xylose-fermenting recombinant yeast strain constructed through a modified genome shuffling method. *Biotechnol. Biofuels* 5 (1), 46. doi:10.1186/1754-6834-5-46

Zhao, H., Zhang, R., Yan, X., and Fan, K. (2021). Superoxide dismutase nanozymes: An emerging star for anti-oxidation. *J. Mater Chem. B* 9 (35), 6939–6957. doi:10.1039/d1tb00720c

Zheng, Y. L., Han, J. M., Wang, B. Y., Hu, X. Y., Li, R. X., Shen, W., et al. (2019). Characterization and repurposing of the endogenous Type I-F CRISPR-Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res.* 47 (21), 11461–11475. doi:10.1093/nar/gkz940

Zhou, Y., Lv, H., Li, H., Li, J., Yan, Y., Liu, F., et al. (2021). Nitroreductase Increases menadione-mediated oxidative stress in *Aspergillus nidulans*. *Appl. Environ. Microbiol.* 87 (24), e0175821. doi:10.1128/aem.01758-21

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fbioe.2022.1109929

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Jinhui Feng,
Tianjin Institute of Industrial Biotechnology (CAS), China
Shihui Yang,
Hubei University, China

*CORRESPONDENCE
Xixi Zhao,
xx.zhao@siat.ac.cn
Xiaozhou Luo,
xz.luo@siat.ac.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 28 November 2022
ACCEPTED 23 December 2022
PUBLISHED 10 January 2023

CITATION
Li F, Deng H, Zhong B, Ruan B, Zhao X and Luo X (2023), Identification of an indole biodegradation gene cluster from *Providencia rettgeri* and its contribution in selectively biosynthesizing Tyrian purple. *Front. Bioeng. Biotechnol.* 10:1109929. doi: 10.3389/fbioe.2022.1109929

Identification of an indole biodegradation gene cluster from *Providencia rettgeri* and its contribution in selectively biosynthesizing Tyrian purple

Feifei Li[1,2,3,4], Huaxiang Deng[2,3,4], Biming Zhong[2,3,4], Banlai Ruan[2,3,4], Xixi Zhao[2,3,4]* and Xiaozhou Luo[1,2,3,4]*

[1]School of Life Sciences, Inner Mongolia University, Hohhot, China, [2]Shenzhen Key Laboratory for the Intelligent Microbial Manufacturing of Medicines, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [3]CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [4]Center for Synthetic Biochemistry, Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Tyrian purple, mainly composed of 6, 6′-dibromoindigo, is a precious dye extracted from sea snails. In this study, we found Tyrian purple can be selectively produced by a bacterial strain GS-2 when fed with 6-bromotryptophan in the presence of tryptophan. This GS-2 strain was then identified as *Providencia rettgeri* based on bacterial genome sequencing analysis. An indole degradation gene cluster for indole metabolism was identified from this GS-2 strain. The heterologous expression of the indole degradation gene cluster in *Escherichia coli* BL21 (DE3) and *in vitro* enzymatic reaction demonstrated that the indole biodegradation gene cluster may contribute to selectively biosynthesizing Tyrian purple. To further explore the underlying mechanism of the selectivity, we explored the intermediates in this indole biodegradation pathway using liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry (LC-ESI-QTOF-MS/MS), which indicated that the indole biodegradation pathway in *Providencia rettgeri* is the catechol pathway. Interestingly, the monooxygenase *GS*-C co-expressed with its corresponding reductase *GS*-D in the cluster has better activity for the biosynthesis of Tyrian purple compared with the previously reported monooxygenase from *Methylophaga aminisulfidivorans* (MaFMO) or *Streptomyces cattleya* cytochrome P450 enzyme (CYP102G4). This is the first study to show the existence of an indole biodegradation pathway in *Providencia rettgeri*, and the indole biodegradation gene cluster can contribute to the selective production of Tyrian purple.

KEYWORDS

Tyrian purple, *Providencia rettgeri*, indole biodegradation gene cluster, selective Tyrian purple producing, monooxygenase

1 Introduction

Tyrian purple, known as royal purple, and mainly composed of 6, 6′-dibromoindigo, is an ancient dye extracted from the murex shellfish (Ngangbam et al., 2015; Lee et al., 2021). Tyrian purple has a range of striking purple to red, color-fast and resistance to fading, and also has a promising application in dye-sensitized solar cells, functional polymers, and conductive materials (Głowacki et al., 2012; Głowacki et al., 2013; Guo et al., 2015; Kim et al., 2018;

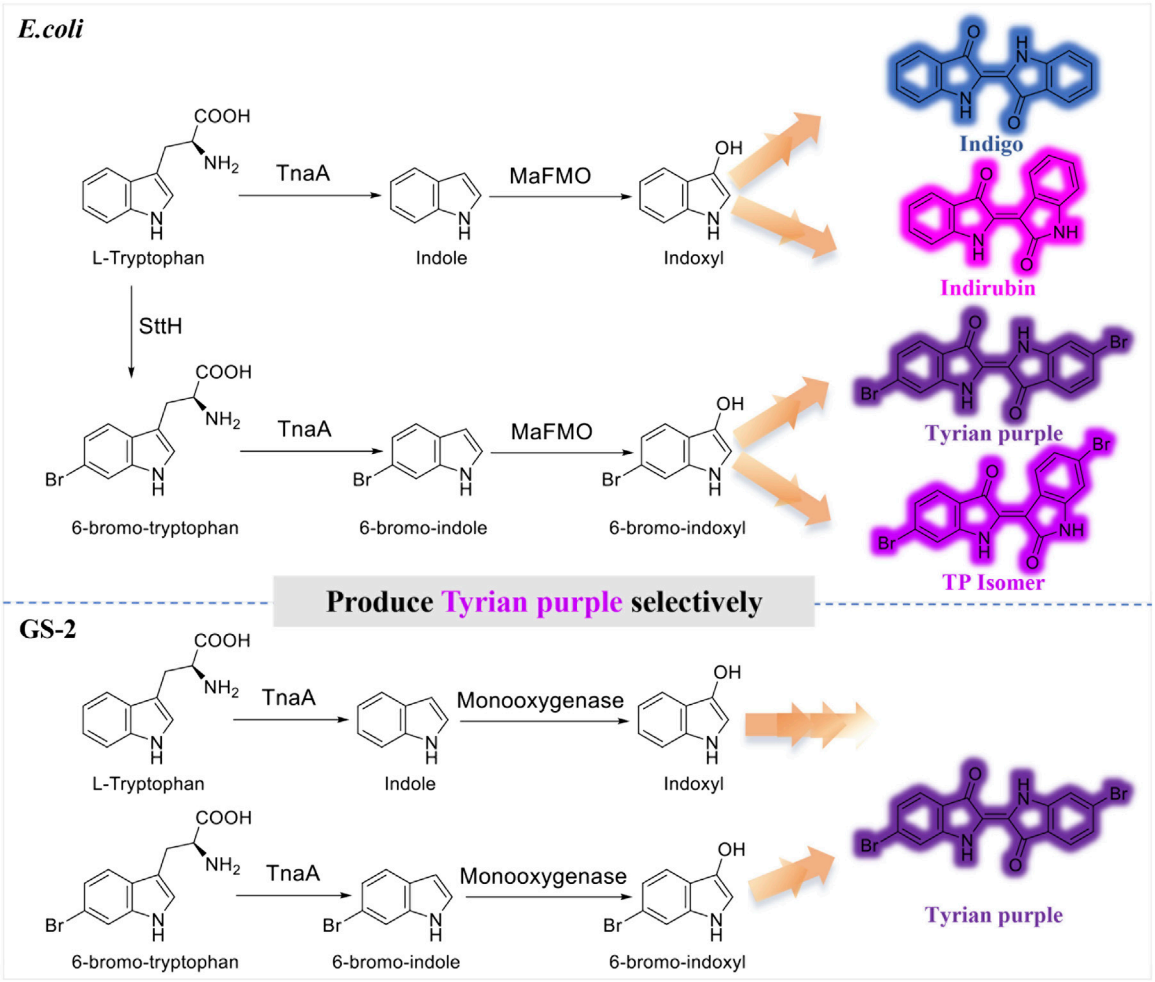


FIGURE 1
Production of indigo, indirubin, Tyrian purple, and TP isomer from tryptophan or 6-bromo-tryptophan: ***E. coli***, expressing tryptophan 6-halogenase SttH, and flavin-containing monooxygenase MaFMO from *Methylophaga aminisulfidivorans* in *E. coli*. **GS-2**, in the bacterial strain GS-2.

Schnepel et al., 2021). It is very difficult to obtain the dye in large quantities from natural sources as it requires euthanizing 12,000 snails per 1.4 g of dye (Mcgovern and Michel, 1990). In addition, due to the difficulty in chemical or biological synthesis, there is still no method for its industrial production (Wolk and Frimer, 2010). Biocatalysis has emerged as an alternative for sustainable synthesis of Tyrian purple from a natural substrate through microbial fermentation, but the selectivity issue in enzymatic tryptophan and bromotryptophan degradation becomes an obstacle for large-scale biosynthesis. Recently, Lee et al. presented an alternative 6, 6′-dibromoindigo production strategy in *E. coli* using tryptophan 6-halogenase from *Streptomyces toxytricini* (SttH), tryptophanase from *E. coli* (TnaA), and monooxygenase MaFMO.

It is often not feasible to obtain pure Tyrian purple by biosynthesis. As TnaA is effective enough to convert tryptophan or its halogenated derivative into the corresponding indoles, and MaFMO catalyzes the hydroxylation of indoles (indole or 6-bromoindole) to 2-hydroxyindoles and 3-hydroxyindoles (Kim et al., 2019), the biosynthesis of Tyrian purple using three enzymes in *E. coli* would produce indigo and indirubin (isomer of indigo), Tyrian purple, and 6, 6′-dibromoindirubin (TP isomer, Tyrian purple isomer), simultaneously (Figure 1), that is, TnaA competes with SttH for tryptophan, both TnaA and MaFMO are lack of selectivity that lead to the production of many byproducts, which are the limitations of Tyrian purple biosynthesis. To overcome this TnaA competes with SttH for tryptophan issue, Lee et al. introduced a consecutive two-cell reaction system to overproduce regiospecifically brominated precursors of 6, 6′-dibromoindigo by spatiotemporal separation of bromination and bromotryptophan degradation (Lee et al., 2021). These approaches led to 315.0 mg L^{-1} 6, 6′-dibromoindigo production from 2.5 mM tryptophan. However, the final product is impure and the two-cell reaction process is still too inefficient and uneconomical to be feasible on an industrial scale for the biological synthesis of Tyrian purple.

Nevertheless, indoles are a typical class of N-heterocyclic aromatic pollutants and are widespread in our daily products and natural environment. In microbial communities, more than 85 bacterial species, including *Escherichia coli, Vibrio cholera*, and *Providencia rettgeri*, can catalyze tryptophan to indole by tryptophanase (TnaA) (O'hara et al., 2000; Lee and Lee, 2010; Ngangbam et al., 2015). However, indole and its derivatives at high concentrations are mutagens and carcinogens that exhibit toxic activity on

microorganisms and animals (Lin et al., 2015; Tomberlin et al., 2017; Li et al., 2021). To defend against the toxicity of indole, many bacteria have established enzymatic detoxification systems, which is the oxidation of indole to insoluble non-toxic indigoid pigments or use the biodegradation mechanisms (Doukyu and Aono, 1997; Bhushan et al., 2000; Kim et al., 2003; Fukuoka et al., 2015; Zhang et al., 2018). A number of indole-degrading bacterial microorganisms and bacterial consortia were reported previously. *Acinetobacter*, *Alcaligenes*, *Burkholderia*, *Pseudomonas*, and *Cupriavidus* are the most extensively investigated indole-degrading bacterial genera (Yin et al., 2005; Kim et al., 2016). There are several reports on the identification of the indole biodegradation gene cluster and using it for the biological production of indigo in *E. coli*, but there are few reports on its application in the biosynthesis of indigoid dyes. Lin et al. identified an indigo-producing oxygenase *iacA* in *Acinetobacter baumannii* and that the iac gene cluster of *A. baumannii* is involved in indole 3-acetic acid degradation (Lin et al., 2012). Qu et al. isolated and unveiled the biotransformation mechanism of indole in *Cupriavidus sp*. strain SHE (Qu et al., 2015a; Qu et al., 2017). Later, they isolated indole-degrading bacterium *Burkholderia sp*. IDO3 and found an *iif2* gene cluster for indole degradation and indigo production (Ma et al., 2019a; Ma et al., 2019b).

Although the indole degradation pathway has been studied for almost a century, the indole biodegradation gene cluster that can contribute to the selective Tyrian purple production has not been reported so far. In this study, we found Tyrian purple can be selectively produced in a *Providencia rettgeri* bacterial strain GS-2 from a laboratory environment. Then, we isolated and obtained the complete genomic sequence of GS-2 and identified an indole degradation gene cluster for indole metabolism in GS-2. Heterologous expression of these genes is from the indole degradation gene cluster in *Escherichia coli* BL21 (DE3). The *in vivo* and *in vitro* reaction results showed that the indole biodegradation gene cluster may contribute to the selective Tyrian purple production. Compared with the most reported monooxygenase MaFMO and cytochrome P450 enzyme CYP102G4, the Tyrian purple-producing enzyme *GS*-C co-expressed with its cofactor *GS*-D in the indole biodegradation gene cluster have the best activity in the production of Tyrian purple in *E. coli*. At last, we tried to unlock the biotransformation and degradation mechanism of selective Tyrian purple producing in GS-2. Unveiling the selective Tyrian purple production will open an avenue to promote the biosynthesis of Tyrian purple.

2 Materials and methods

2.1 Bacterial strains, chemicals, and standard techniques

Escherichia coli DH5α and BL21 (DE3) strains were used as cloning and protein expression hosts, respectively. *E. coli* strains carrying plasmids were cultivated in a Luria-Bertani (LB) medium supplemented with antibiotic (50 μg/mL kanamycin). An LB medium was purchased from Huankai Microbial (China). M9 media salts were purchased from Sangon Biotech (China). Plasmid DNA was isolated using a Tiangen plasmid miniprep kit (TIANGEN, Ltd. China). 6-Bromo-tryptophan and 6-chloro-tryptophan were purchased from GL Biochem (China). Tryptophan and NADH were purchased from Sigma (United States). Indole, 6-bromo-indole, indigo, FAD, and isatin were purchased from Sigma (United States), Bidepharm (China), and Yuanye Biotech (China), respectively. Indirubin was purchased from TCI (Japan). Chemically synthesized 6, 6′-dibromoindigo was synthesized from Abace Biotech (China). Gene and oligomer synthesis and sequencing were carried out in BGI (China), Sangon (China), and Rui Biotech (China), respectively. Enzymes involved in restriction reaction, ligation, and PCR were purchased from New England Biolabs (United States), Vazyme (China), and TAKARA (Japan), respectively. All other chemicals used in this study were of analytical grade. All media and reagent solutions were prepared with Milli-Q water (Merck Millipore).

2.2 Genome sequencing and annotation

The bacterial strain designated as GS-2 was isolated from the LB agar plates supplemented with 6-bromo-tryptophan (containing purple colonies). To obtain a partial genome sequence, total DNA from GS-2 was isolated using the TIANamp Bacteria DNA Kit (TIANGEN, China). The genome of GS-2 was sequenced with 15775578 reads of an average length of 350 bp by using an Illumina HiSeq 4000 system (Illumina, San Diego, CA, United States) at the Beijing Genomics Institute (Shenzhen, China). Raw reads of low quality from paired-end sequencing were discarded. The sequenced reads were assembled using SOAP *de novo* v1.05 software. Gene prediction was performed on the GS-2 genome assembly by Glimmer3 (http://www.cbcb.umd.edu/software/glimmer/) with hidden Markov models. tRNA, rRNA, and sRNAs recognition made use of tRNAscan-SE (Lowe and Eddy, 1997), RNAmmer, and the Rfam database. The best hit abstracted using a BLAST alignment tool for function annotation. Seven databases which are KEGG (Kyoto Encyclopedia of Genes and Genomes), COG (Cluster of Orthologous Groups), NR (non-redundant protein database), Swiss-Prot, and GO (Gene Ontology), TrEMBL, and EggNOG are used for general function annotation.

2.3 Identification of an indole biodegradation gene cluster

According to the reported microbial indigo-forming enzymes or indole biodegradation gene cluster and the draft genome sequencing of GS-2 and bioinformatics analysis of the genome of GS-2 using NCBI's blast$^+$ algorithm, blastn algorithm, and the Conserved Domain Database (CCD), we use the reported microbial indigo-forming enzymes or indole biodegradation gene cluster as the user's query to the database of the GS-2's draft genome sequences. The prevalence of indole biodegradation gene cluster as well as genes of indole oxygenase among available microbial genomes was analyzed by using the NCBI's blast$^+$ algorithm, too. Based on the blastn algorithm and the gene annotation of *Providencia heimbachae* ATCC 35613 in NCBI, seven genes of the *GS* (A-R) encoded putative enzymes, as described by conserved domain analysis.

2.4 Heterologous expression of indole oxygenase genes and indole biodegradation gene cluster

2.4.1 Plasmid and strain construction

Genes *GS* (A-E) from genomic DNA of strain GS-2 were amplified using 2×Phanta® Master Mix (Vazyme, China). The primers for

TABLE 1 List of oligos used in this study.

Primer	Sequence
pET28a-fwd	CACCACCACCACCACCACTGAGATC
pET28a-Rev	CATGAATTCGGATCCGCGACCC
GS-C-fwd	GTCGCGGATCCGAATTCATGCGTCGTATTGCTATCG
GS-C-rev	CAGTGGTGGTGGTGGTGGTGTTAGGCTCGAGCTTCTTC
GS-C + D-fwd	AACTTTAAGAAGGAGATATACCATGATGTCTATGATTGCAGAGGATAC
GS-C + D-rev	CATCATGGTATATCTCCTTCTTAAAGTTAATTAGGCTCGAGCTTCTTCCATC
GS-D-rev	CAGTGGTGGTGGTGGTGGTGCTAAAGGGCTTCATGTAAAG
GS-B-fwd	CAAATGGGTCGCGGATCCGAATTCATGAGAGTAAAAAAGAGGATGCCGAAG
GS-B + C + D + A-rev	CATGGTATATCTCCTTCTTAAAGTTAATTATTCAGCAAGACGGGGAATTG
GS-B + C + D + A-fwd	AACTTTAAGAAGGAGATATACCATGATGAACCAGCCAGCATTGATG
GS-A-rev	GATCTCAGTGGTGGTGGTGGTGGTGTCACCCCACCAGCGCTAG
GS-O-fwd	GTCGCGGATCCGAATTCATGACAACACACTCTCACCC
GS-O-rev	CAGTGGTGGTGGTGGTGGTGTTATGTCCTACGTTTGAC
GS-O + E-rev	TGGTATATCTCCTTCTTAAAGTTAATTATGTCCTACGTTTGACTGCC
GS-E-fwd	TTAACTTTAAGAAGGAGATATACCATGATGGCAGGAAATATGATGAAAAC
GS-E-rev	GATCTCAGTGGTGGTGGTGGTGGTGTTAGAATGGAATGGCCGC
GS-B + C + D + A + O + E-fwd	CAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGATGACAACACACTCTCACCCTCAAG
GS-B + C + D + A + O + E-rev	CAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTATCACCCCACCAGCGCTAG
TnaA-fwd	ATGGAAAACTTTAAACATCTCCC
TnaA-rev	AACTTCTTTAAGTTTTGCGGTG
MaFMO-fwd	GAAGGAGATATACCATGATGGCAACTCGTATTGCGATAC
MaFMO-rev	ATCTCCTTCTTAAAGTTAATTAAGCTTCTTTAGCCACAGG
CYP102G4-fwd	TCGAGTGCGGCCGCAAGCTTTCACCCGGCGGCGTACAC
CYP102G4-rev	CGGTGGCAGCAGCCTAGGTTAACCTGCTGCATAAACATC

amplification are provided in Table 1. The PCR products were purified and ligated with linearized pET28a (+) plasmid using a Gibson One Step Cloning Kit (New England Biolabs, United States). Tryptophanase gene (TnaA) from *E. coli* was amplified from the genomic DNA of *E. coli* BL21 (DE3) using primer TnaA_fwd and TnaA_rev, and it was cloned into pCDF, yielding pCDF:tnaA. To transform host cells with the fusion constructs, the Gibson reaction mixture (5 μL) was added to chemical competent *E. coli* DH5α cells, and a heat shock (42°C) was applied for 60 s. After overnight growth on an LB agar plate with kanamycin, colonies were picked and grown in LB, and then the plasmids were isolated and sent for sequencing (Rui Biotech, China) to confirm the correct ligation of the genes. Then, the recombinant plasmid was successfully expressed in strain *E. coli* BL21 (DE3).

2.4.2 Growth curve, growth conditions, and protein expressions

Overnight cultured GS-2 and *E. coli* were inoculated into a fresh medium, and the OD_{600} of bacteria was measured using a smart microplate reader (Infinite® 200 Pro, Tecan, Switzerland) every 30 min. The experiment was repeated three times. Expression plasmids with N-terminally encoded 6×His tags were transformed into *E. coli* BL21 (DE3) for protein expression. A single colony was inoculated into an LB medium with corresponding antibiotics and grown overnight at 37°C. Then, 30 μL of the seed cultures were inoculated into 3 mL of an LB media, and they were grown at 37°C until the cell density reached an optical density (OD600) of roughly 0.4–0.6. Expressions of proteins were induced by the addition of 0.1 mM isopropyl-β-d-thiogalactoside (IPTG) with subsequent overnight incubation at 30°C. Expressions of TnaA + CYP102G4 were induced by 0.1 mM IPTG and 0.25 mM ALA (5-aminolevulinic acid) with overnight incubation at 30°C.

2.4.3 Purification of proteins and enzyme assays

Expressions of *GS*-C, *GS*-D, *GS*-B, *GS*-A, GDH, and MaFMO were induced by 0.1 mM IPTG in 1 L of M9 media with overnight incubation at 18°C. The cells were collected by centrifugation at 9,000 rpm for 10 min and washed with 50 mM Tris-HCl (pH 7.5) buffer solution. The cells were resuspended in Tris buffer solution containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 10% (v/v)

glycerol. The cell soup was disrupted by ultrasonication in ice-chilled water for 40 min (2 s turn on and 2 s turn off). The soluble fractions were collected by centrifugation of the cell lysate at 12,000 rpm for 60 min at 4°C. The soluble fractions of the proteins were purified by Ni-NTA His-tag protein purification. The total and soluble fraction of lysates and the elution fractions were loaded into 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) at 250 V for 32 min. The elution fractions were concentrated to 3 mL using an Amicon 50 mL 10 or 30 kDa cutoff centrifugal filter. Protein concentration was determined by using the A280 feature of a Thermo Scientific 2000 Nanodrop.

The molecular weight of the purified proteins was detected by MALDI-Tof MS as follows: 2 μL of the purified protein was transferred onto an MTP 384 polished steel target (Bruker Daltonics, Billerica, MA) using a 2.5 μL pipette tip; 1 μL of α-cyano-4-hydroxycinnamic acid (CHCA) solution (acetonitrile: H_2O: trifluoroacetic acid (TFA) = 50:47.5:2.5, v/v) was then spotted. MALDI-ToF mass spectra were acquired using a Bruker Autoflex MALDI-ToF/ToF mass spectrometer (Bruker Daltonics) equipped with a frequency tripled Nd:YAG solid state laser (λ = 355 nm). Mass spectrometer calibration was performed using Peptide Calibration Standard Kit II (Bruker Daltonics). Spectral acquisition was performed in positive reflection mode with pulsed ion extraction and a mass range of 10,000–70,000 Da. The laser footprint was set to "Ultra" at a ~100 μm diameter, and 500–1,000 laser shots were fired at 1,000 Hz. Mass spectra were smoothed, baseline-corrected, and analyzed in FlexAnalysis 3 (Bruker Daltonics).

2.4.4 *In vivo* reactions

For *in vivo* substrate reaction, 1 mM tryptophan or 1 mM 6-bromo-tryptophan, 1 mM indole, or 0.25 mM 6-bromo-indole were added into the overnight induced culture, respectively. For the identification of the monooxygenase selectivity, feeding 1 mM tryptophan and 1 mM 6-bromo-tryptophan or 1 mM indole and 0.25 mM 6-bromo-indole simultaneously. Indigo/Tyrian purple concentration was determined by LC-MS analysis described as follows.

2.4.5 Enzyme kinetics and *in vitro* reactions

For the enzyme kinetics reaction, indole oxidation activity of *GS*-C and *GS*-D were determined as a function of indigo/Tyrian purple formation in the reaction mixture, respectively. Initial indole oxidation rather than spontaneous dimerization was assumed to be a rate-limiting step. A typical reaction mixture contained 50 mM Tris-HCl, pH 8.0, 200 μM FAD, 250 μM NADH, and various concentrations of indole (10–1,000 μM) or 6-Br-indole (10–1,000 μM) and 300 μg of *GS*-C as well as 300 μg of *GS*-D were used for the analysis of the enzyme kinetics. A typical reaction mixture contained 50 mM Tris-HCl, pH 7.5, 250 μM NADH, 0.6% glucose, and various concentrations of indole (10–1,000 μM) or 6-Br-indole (10–1,000 μM) and 300 μg of MaFMO as well as 5 U/mL of GDH were used to analyze the enzyme kinetics. Reaction mixtures were incubated at 30°C for 120–240 min. One unit of enzyme activity was defined as the amount catalyzing the formation of 1 μmol of indigo/Tyrian purple per minute.

Flavin reductase activity of the purified *GS*-D protein was determined from the decrease of the absorbance at 340 nm due to the oxidation of NADH, using a microplate reader and was performed at 30°C. A total reaction volume of 0.2 mL contained 50 mM Tris-HCl, pH 8.0, 250 μM NADH, and 100 μM FAD. The reactions were initiated by adding 300 μg of *GS*-D. One unit (U) of enzyme activity was defined as the amount of enzyme catalyzing the oxidation of 1 μmol of NADH per minute.

For *in vitro* substrate reaction, indole oxidation activity of *GS*-A, *GS*-B, *GS*-C, and *GS*-D were determined as a function of indigo/Tyrian purple formation in the reaction mixture, respectively. A typical reaction mixture contained 50 mM Tris-HCl, pH 8.0, 200 μM FAD, 250 μM NADH, and 1,000 μM indole or 1,000 μM 6-Br-indole and 100 μg of *GS* proteins were used for the *in vitro* substrate reaction. Reaction mixtures were incubated at 30°C for 150 min. Indigo/Tyrian purple concentration was determined by LC-MS analysis described as follows.

2.5 Identification of metabolite

For qualitative and quantitative analyses of indigo/Tyrian purple, 0.1 mL of the reaction mixtures were centrifuged at 12,000 rpm, 10 min, and the supernatant was removed. The pellet was suspended in 1 mL of DMSO followed by vigorous vortexing. The mixtures were centrifuged at 12,000 rpm, 10 min, and the supernatant was filtered through 0.22 μm filters for LC-MS analysis. For LC-MS analysis of the solid plate of *E. coli* strain expressing different genes by *in vivo* reactions, the pellet was resuspended in 1 mL DMSO and the supernatant was sent for LC-MS analysis after centrifugation. LC connected with tandem mass spectrometry (Agilent 6470 QQQ) using the C18 reversed-phase (4.6 × 100 mm, 2.7-Micron). LC-MS data were collected using the Agilent MassHunter Workstation. The condition for LC gradient wash was as follows: a gradient of H_2O+ 10 mM ammonium acetate (solvent A) and acetonitrile (solvent B) using the following method: 60%–50% B for 16–40 min, at a flow rate of 0.5 mL min^{-1}. The mass spectrometry conditions were as follows: column temperature 40°C; electrospray ionization in negative mode; capillary voltage 3.8 kV; vaporizer temperature 350°C; capillary temperature 320°C; sheath gas pressure 30 psi; auxiliary gas pressure 10 psi; and SIM event 2, mass for the analysis of indigo and indirubin is 261. Mass for the analysis of 6, 6′-dibromoindigo is Tyrian purple, TP; m/z 417, 419, and 421 and 6, 6′-dibromoindirubin is TP isomer; m/z 417, 419, and 421.

For liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry (LC-ESI-QTOF-MS/MS) analysis, the sample extraction was as follows: 0.2 mL of the sample from each reaction was collected in a new 1.5 mL centrifuge tube and freeze dried for 3–5 h. The solid was then dissolved in 1 mL ethanol with 0.1% formic acid and treated with ultrasound for 1.5 h. Each sample was centrifuged (12,000 rpm, 10 min) to take the supernatant, which was filtered through 0.22 μm filters for analysis. The condition for LC gradient wash was as follows: a gradient of H_2O+ 0.1% formic acid (solvent A) and acetonitrile +0.1% formic acid (solvent B) using the following method: 2% B for 1 min, 2%–40% B for 4 min, 40%–70% B for 5 min, 70%–95% B for 3 min, and 95% B for 4 min and re-equilibration 2% B for 3 min. A volume of 5 μL was injected for each standard or sample and the flow rate was set at 0.3 mL/min. Nitrogen gas nebulization was set at 45 psi with a flow rate of 10 L/min at 350°C, and the sheath gas was set at 12 L/min at 350°C. The capillary and nozzle voltage were set at 4 kV and 1,500 V, respectively. A complete mass scan ranging from m/z 50 to 1,300 was used, and MS/MS analyses were carried out in automatic mode with collision

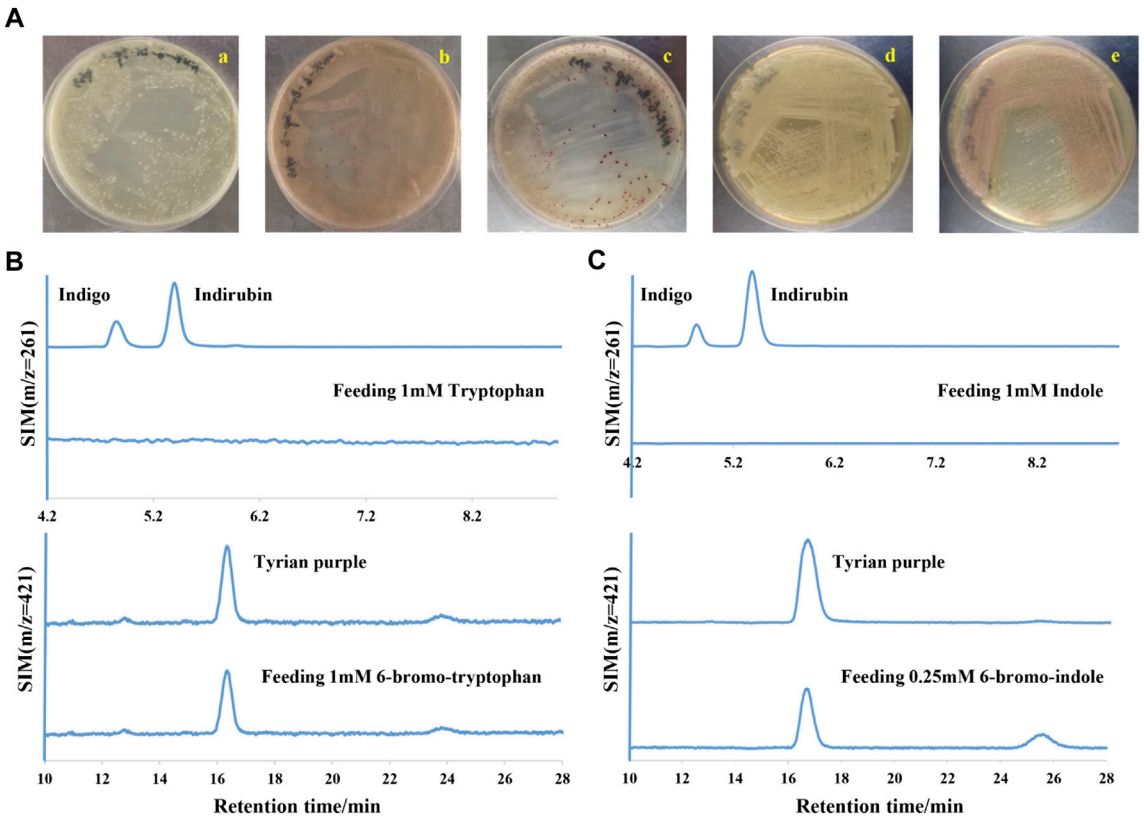


FIGURE 2
Tyrian purple can be selectively produced in GS-2. **(A)** GS-2 grown on the LB agar plates supplemented with different substrates: a. 1 mM Tryptophan; b. 1 mM 6-chloro-tryptophan; c. 1 mM 6-bromo-tryptophan; d. 1 mM Indole; and e. 0.25 mM 6-bromo-indole. **(B)** Extracted ion chromatograms for standards (the upper) of indigo (m/z = 261), indirubin (m/z = 261), Tyrian purple (m/z = 421), and colonies (the lower) were collected from GS-2 grown on the LB agar plates supplemented with tryptophan or 6-bromo-tryptophan, respectively. **(C)** Extracted ion chromatograms for the standards (the upper) of indigo (m/z = 261), indirubin (m/z = 261), Tyrian purple (m/z = 421), and colonies (the lower) were collected from GS-2 grown on the LB agar plates supplemented with indole or 6-bromo-indole, respectively.

energy (10, 20, and 40 eV) for fragmentation. Peak identification was performed in both positive and negative modes while the instrument control, data acquisition, and processing were performed using MassHunter Workstation software (Qualitative Analysis, version 10.0) (Agilent Technologies, Santa Clara, CA, United States).

2.6 Indole, 6-Br-indole, Tyrian purple, indirubin, and indigo toxicity test

To test the effects of different concentrations of indole, 6-Br-indole, Tyrian purple, indirubin, and indigo on the growth of *E. coli* BL21 (DE3), 3 mL of overnight LB culture media was collected and centrifuged. The supernatant was discarded, and the pellet was resuspended in 1 mL of LB. This aliquot was then transferred to a flask containing 70 mL of LB and incubated. Thus, after adding varying amounts of indole, 6-Br-indole (dissolved in absolute ethanol), Tyrian purple, indirubin, and indigo (dissolved in dimethylsulfoxide), the culture was incubated at 37°C, 220 rpm for 17 h. The optical density at 600 nm (OD600) was used to determine the growth of the bacteria under the various conditions. The experiments were performed in triplicate.

3 Results

3.1 Tyrian purple can be selectively produced in a bacterial strain

In a laboratory environment, we accidentally found some purple colonies as contaminants on Luria-Bertani (LB) agar plates supplemented with 1 mM 6-bromo-tryptophan (6-Br-Trp) or 1 mM 6-chloro-tryptophan (6-Cl-Trp) (Figure 2A). Based on the previous study on biological synthesis of 6, 6′-dibromoindigo (Tyrian purple) from tryptophan (Trp), 6-Br-Trp is an intermediate in the Tyrian purple biosynthesis pathway (Lee et al., 2021). Whether these purple colonies were caused by the presence of Tyrian purple? To confirm our hypothesis, the purple colonies were collected and dissolved into dimethylsulfoxide (DMSO) and analyzed by high-performance liquid chromatography mass spectrometry (LC-MS). The results showed that the purple pigments were Tyrian purple (Figure 2B). To validate the Tyrian purple biosynthesis pathway from 6-Br-Trp, 6-bromo-indole (6-Br-indole) was evaluated through substrate reaction. Notably, this strain produced Tyrian purple when grown on an LB agar plate supplemented with 0.25 mM 6-Br-indole (Figure 2C). Halo-tryptophan is tryptophan analogs, if a

TABLE 2 Bacteria genome sequencing result of *Providencia rettgeri* GS-2.

Sample name	Tax ID	Organism	Cover_Len (bp)	Scaffolds_Len	Coverage (%)	Genomics (%)	Scaffold_N
Mut.GS-2	587	*Providencia rettgeri*	4266344	4658124	91.59	95.55	50
	546	*Citrobacter freundii*	153186	153186	100.00	3.14	1
	1157951	*Providencia stuartii*	16099	24891	64.68	0.51	2
	6239	*Caenorhabditis elegans*	13263	13275	99.91	0.27	11
	333962	*Providencia heimbachae*	5473	12461	43.92	0.26	2

Tax ID: species tax id, Organism: species name, Cover_aLen: Alignment coverage length, Scaffolds_Len: alignment to the length of the sequenced strain scaffold on this species, Coverage: alignment length as a percentage of the scaffold length, Genomics: scaffold length as a percentage of the sequence length of the assembly genome of the sequenced strain, Scaffold_Num: the number of scaffolds aligned to the species.

strain can utilize halo-tryptophan to produce purple pigments, it should also mediate the indigo and indirubin production from tryptophan (Namgung et al., 2019; Schnepel et al., 2021) (Figure 1). However, we verified there was neither indigo nor indirubin production in the colonies from LB agar plates supplemented with the tryptophan or indole (Figure 2). According to the results, we supposed this strain can produce Tyrian purple selectively from Trp/6-X-Trp (X = Cl, Br) or indole/6-X- indole (X = Cl, Br).

3.2 Characterization of strain GS-2

The growth curve of GS-2 based on OD is shown in Supplementary Figure 1A. The curve shows that GS-2 reached the stationary phase at about 10 h. To further explore the underlying mechanism for the selectivity, the bacterial strain found to be a selectively Tyrian purple producer was isolated and designated as GS-2. The bacterial genome sequencing analysis predicted that GS-2 is *Providencia rettgeri* with tax number 587. The genome size of GS-2 is 4,874,999 bp. The 4,658,124 bp obtained reads were *de novo* assembled into 79 scaffolds, with a 95.55% of the assembly genome of the sequenced strain (Table 2). The alignment length is 4,266,344 bp, with a 91.59% of the scaffold length. The genome of strain GS-2 had an overall GC content of 41.52%. A total of 4,565 genes, including 137 RNAs, were predicted in the genome. The genome sequencing result of GS-2 have been successfully submitted to NCBI (SRA accession number: JAPQLO000000000), and the genomic information of strain GS-2 facilitates the molecular mechanism and bioremediation study of *Providencia rettgeri*.

Providencia is a ubiquitous Gram-negative bacterium in the family of *Enterobacteriaceae*, which are considered as opportunistic pathogens (Galac and Lazzaro, 2012; Sadauskas et al., 2017). *Providencia rettgeri* is an indole-positive bacterium, which can produce indole by TnaA from L-tryptophan. According to several reports on the biological production of indigo itself in *E. coli*, various colored indigoid dyes can be generated by feeding appropriate halogenated indoles as a substrate to oxygenases, naphthalene dioxygenase, and toluene dioxygenase (Qu et al., 2012; Zhang et al., 2013; Fukuoka et al., 2015; Heine et al., 2019; Choi, 2020; Mendoza-Avila et al., 2020; Lee et al., 2021). When the GS-2 strain was grown on LB agar plates supplemented with Trp, 6-Br-Trp or indole, and 6-Br-indole, Tyrian purple was the only indigoid product (Figure 2). These results indicated that there may exist a selective oxygenase (only oxidize halogenated indole) in strain GS-2. Otherwise, there may be a selective indole biodegradation gene cluster (degrade the indole, but not 6-Br-indole) in strain GS-2.

3.3 Bioinformatics analysis of the genes involved in indole degradation and oxidation

Bioinformatics analysis of the genome of GS-2 by the NCBI's blast$^+$ algorithm with the reported microbial indigo-forming enzymes or indole biodegradation gene cluster allowed the identification of a set of genes (here designated *GS* (A-R)) (Table 3) for indole degradation and oxidation (Ma et al., 2018; Fabara and Fraaije, 2020). These genes *GS* (A-R), which were located in a 7-kb genomic fragment of GS-2, were further analyzed based on the NCBI's blastn algorithm and the Conserved Domain Database (CCD) with other *Providencia* species. According to the prediction of *Providencia heimbachae* ATCC 35613, these genes *GS* (A-R) involved in the indole degradation and oxidation were described as the conserved domain predication, such as *GS*-A encodes a cyclase family protein, arylformamidase activity. *GS*-B encodes a short-chain dehydrogenase/reductase and oxidoreductase activity. *GS*-C encodes a styrene monooxygenase. *GS*-D encodes a flavin reductase, FMN binding, and monooxygenase activity. *GS*-E is a putative MetA-pathway of phenol degradation. *GS*-O encodes oxidoreductase. In addition, a putative AraC transcription regulator, *GS*-R occurred.

To further predict the functions of *GS* (A-R) genes, we analyzed the prevalence of *GS* (A–R) and indigo-forming genes among available microbial genomes (Figure 3). Comparing with those species (*Acinetobacter*, *A. baumannii* and *Cupriavidus*) have the ability of indole degradation and indigo formation (Lin et al., 2012; Sadauskas et al., 2017; Ma et al., 2019a; Ma et al., 2019b), the *GS*-C and *GS*-D genes were found to be a common component among different microbials with high similarity. From previous research, *GS*-C homologues genes were identified as indigo-forming enzymes (monooxygenase activity was detected by oxidizing indole to indigo) in many microbials (O'Connor et al., 1997; Drewlo et al., 2001; Doukyu et al., 2003; Alemayehu et al., 2004; Lu and Mei, 2007; Kwon et al., 2008; Ameria et al., 2015; Qu et al., 2015b; Heine et al., 2019; Fabara and Fraaije, 2020). The *GS*-D gene encoded a flavin reductase, flavin adenine dinucleotide (FAD) cofactor, and NAD(P)H, which supposed that *GS*-C was possibly a cofactor-independent oxygenase that catalyzed the transformation of indole to indigo. *GS*-A and *GS*-B genes were involved in the indole degradation along with *GS*-C and *GS*-D genes. It is worth noting that *GS*-O encodes oxidoreductase appears only in *Providencia*, and the

TABLE 3 Predicted conserved domains and putative functions of proteins encoded in *Providencia rettgeri* GS-2 according the prediction of *Providencia heimbachae* ATCC 35613.

Protein	Gene_id	Length/bp	Identity (%)	E_value	Accession_id	Function and conserved domain
GS-A	Mut.GS.2GL000526	782	90	0	A0A1B7JR90	Arylformamidase activity, cyclase (PF04199)
GS-B	Mut.GS.2GL000529	785	92	$9.3E^{-170}$	A0A1B7JRB6	Short-chain dehydrogenase/reductase, oxidoreductase activity
GS-C	Mut.GS.2GL000528	1244	94	0	A0A1B7JR97	Styrene monooxygenase A putative substrate binding domain
GS-D	Mut.GS.2GL000527	530	89	$6.9E^{-115}$	A0A1B7JR95	Flavin_Reduct (PF01613), FMN binding, and monooxygenase activity
GS-E	Mut.GS.2GL000524	908	91	0	A0A1B7JRB5	Putative MetA-pathway of phenol degradation (PF13557)
GS-R	Mut.GS.2GL000523	899	94	0	A0A1B7JR83	AraC family transcriptional regulator
GS-O	Mut.GS.2GL000525	1613	90	0	A0A1B7JR88	Oxidoreductase activity (GO:0016491)

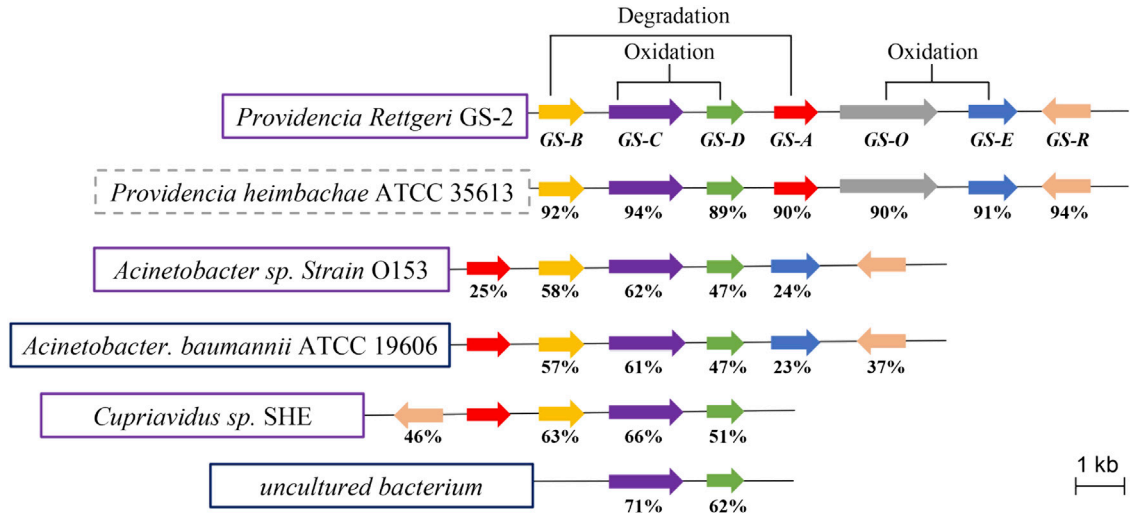


FIGURE 3
Organization and distribution of GS (A-R) genes in different microbial genomes. Genes are represented by arrows (drawn to scale as indicated). Homologous genes are highlighted in the same pattern according to the scheme in the top line for GS-2. Strains and genomic fragments boxed in solid lines indicate reported activity of certain GS-homologous proteins; dashed boxes highlight strains for which indole biodegradation activity was reported at the gene level. Identities (percent) of amino acid sequences between GS proteins of strain GS-2 and homologs are indicated under the corresponding ORFs.

function of *GS*-O has not been identified. Therefore, *GS*-O may be a selective oxygenase in strain GS-2.

3.4 Characterization of the *GS* proteins in *E. coli*

To elucidate the functions of these *GS* proteins, *GS*-A, *GS*-B, *GS*-C, *GS*-D, *GS*-O, and *GS*-E genes were successfully overexpressed in *E. coli* BL21 (DE3) (which possesses TnaA), with an N-terminal His-tag, *GS*-A, *GS*-B, *GS*-C, and *GS*-D were successfully purified, and their activity toward Trp, 6-Br-Trp, indole, and 6-Br-indole was tested, which were monitored by LC-MS. Molecular masses of the purified proteins observed in SDS-PAGE gels and MALDI-TOF mass spectrometry corresponded well to the theoretical masses (32.8 kDa for *GS*-A, 32.9 kDa for *GS*-B, 50 kDa for *GS*-C, and 23.4 kDa for *GS*-D) (Figure 4 A and B).

Bioinformatics analysis of the *GS*-C gene predicted that *GS*-C was possibly an indigo-forming enzyme, so its activities toward indole and 6-Br-indole were evaluated by *in vitro* reactions first. When 1 mM indole was used as the substrate, *GS*-C produced indigo and indirubin (Figure 4C). When 1 mM 6-Br-indole was used as a substrate, Tyrian purple and TP isomer were produced. Though *GS*-C was found to be capable of using indole or 6-Br-indole as substrates, indole and 6-Br-indole were toxic to the growth of *E. coli* (Lin et al., 2015), so Trp and 6-Br-Trp were added to the culture of *GS*-C in *E. coli* BL21 (DE3) (which possesses TnaA, Trp and 6-Br-Trp can be catalyzed to indole and 6-Br-indole, respectively) instead. The transformed product of *GS*-C in *E. coli* BL21 (DE3) supplied with 1 mM Trp substrate was a blue insoluble indigoid pigment, which was identified as indigo and indirubin by LC-MS (Figure 5). The strain supplemented with 1 mM 6-Br-Trp showed the purple pigment, which was identified as Tyrian purple and TP isomer by LC-MS. When feeding 1 mM Trp and 1 mM 6-Br-Trp simultaneously, the transformed product of which was a blue–purple insoluble indigoid pigment (Figure 5A). Hence, we suggested that *GS*-C was functional as oxygenase for Tyrian purple biosynthesis without observable selectivity toward indole or 6-Br-indole.

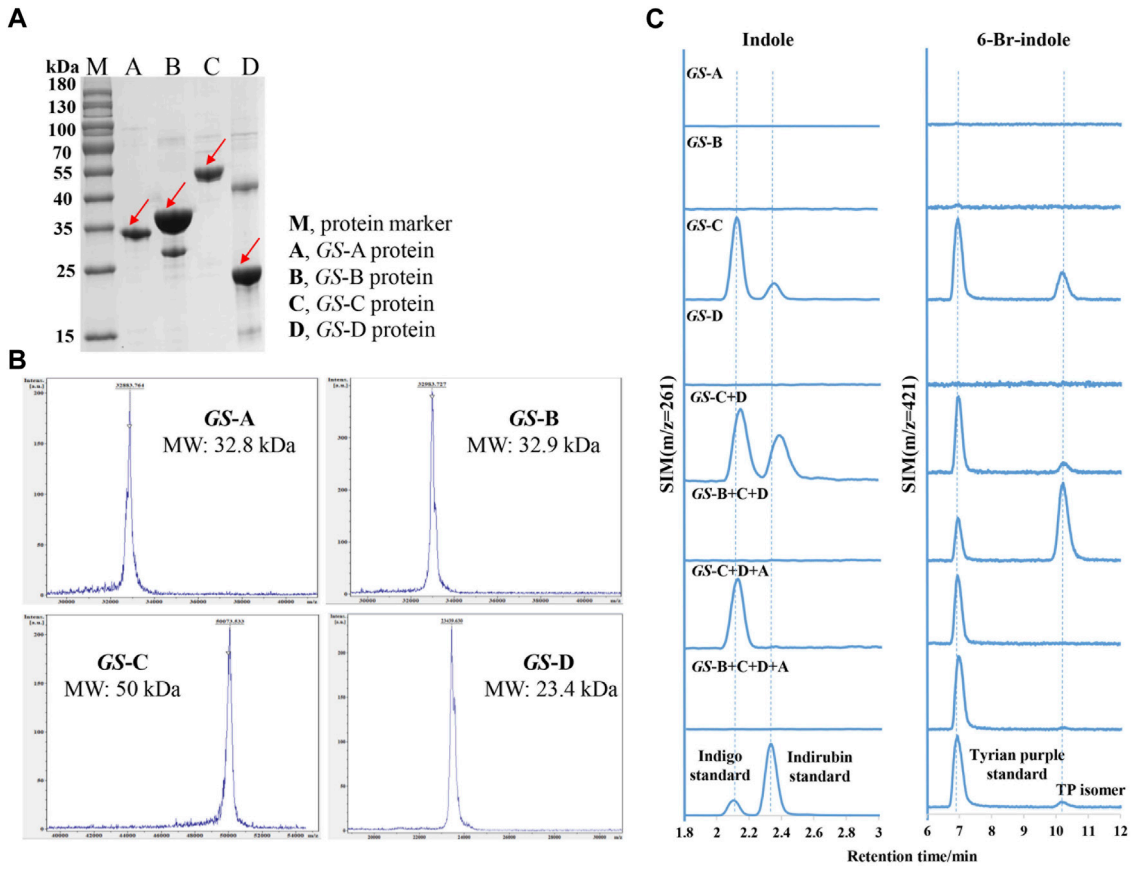


FIGURE 4
Purification and characterization of indole biodegradation gene cluster. **(A)** SDS-PAGE analysis of the purified four *GS* proteins. M, protein marker; A, *GS*-A; B, *GS*-B; C, *GS*-C; D, *GS*-D. **(B)** MALDI-TOF mass spectrometry of the purified four *GS* proteins. **(C)** Extracted ion chromatograms for the standards (the lower) of indigo (m/z = 261), indirubin (m/z = 261), Tyrian purple (m/z = 421), and production of *in vitro* reactions by different *GS* proteins supplemented with indole or 6-bromo-indole, respectively.

The *in vitro* reaction results of *GS*-D showed that there was neither Tyrian purple nor indigo/indirubin, and TP isomer be detected by LC-MS when 1 mM indole or 1 mM 6-Br-indole was used as substrate. Based on the properties of two-component FAD-dependent indole monooxygenases, both protein components are usually encoded next to each other on the genome in a respective gene cluster (Heine et al., 2018). Since *GS*-D was proposed as a cofactor generating enzyme and located near *GS*-C in the genome, we were wondering whether the addition of *GS*-D to *GS*-C strain can improve the activity of *GS*-C, and the *in vivo* reaction results showed that with *GS*-D added, there were more blue/purple pigments than which produced by *GS*-C in *E. coli* (Figure 5A). It is easy to distinguish the differences between *GS*-C and *GS*-C + D by the results of ion chromatograms (Figure 5B). We also characterized the flavin reductase activity of the purified *GS*-D reductase. As shown in Supplementary Figure 2, the decrease of the absorbance at 340 nm due to the oxidation of NADH demonstrated the flavin reductase activity of the *GS*-D reductase. The catalytic results of *GS*-C and *GS*-D are consistent with the homologous enzymes reported previously (Dai et al., 2019; Heine et al., 2019).

To identify whether the *GS*-O is a selectivity oxygenase in the strain GS-2, *GS*-O was expressed in *E. coli* BL21 (DE3) and cultured with Trp and 6-Br-Trp, there was no colony with any observable blue or purple color on the plates (Figure 5A), and none of Tyrian purple, indigo/indirubin, and TP isomer was identified by LC-MS (Figure 5B). These results indicated that the *GS*-O may not work as the Tyrian purple-producing oxygenase in strain GS-2. To further verify the hypotheses that *GS*-O may be capable of using indole as a substrate only in the presence of *GS*-E, *GS*-O and *GS*-E were co-expressed in *E. coli* BL21 (DE3) cultured with Trp and 6-Br-Trp, neither blue pigment nor purple pigment occurred, too (Figure 5A). These results indicated that *GS*-O and *GS*-E were not the selective Tyrian purple-producing oxygenase in strain GS-2.

To identify whether the indole degradation gene cluster is involved in the selective Tyrian purple production in strain GS-2, we added *GS*-B to the reaction mixture (*GS*-C, *GS*-D, FAD, NADH, and indole), the indigo formation was abolished, and TP isomer formation increased. When *GS*-A was added to the reaction mixture (*GS*-C, *GS*-D, FAD, NADH, and 6-Br-indole), TP isomer formation was abolished. When the four *GS* proteins (*GS*-A, *GS*-B, *GS*-C, and *GS*-D) were used in a reaction with indole or 6-Br-indole, Tyrian purple was the only target product. In addition, we co-expressed the genes *GS*-A, *GS*-B with *GS*-C, and *GS*-D in *E. coli* BL21 (DE3) supplemented with Trp and 6-Br-Trp. The recombinant strain produced a lot of yellow colonies on the plate cultured with 1 mM Trp. There were purple insoluble indigoid pigments on the plate cultured with 1 mM 6-Br-Trp. Purple pigment

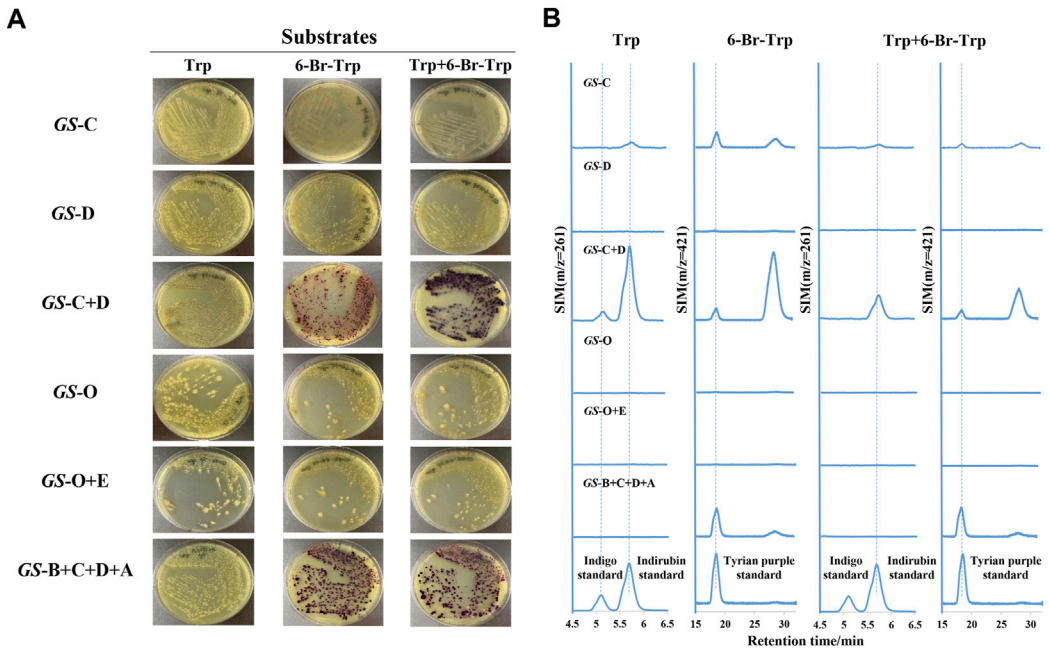


FIGURE 5
Heterologous expression of indole oxygenase genes and indole biodegradation gene cluster in *E. coli* BL21 (DE3). **(A)** Engineered *E. coli* strains with different *GS* proteins were respectively grown on agar plates with tryptophan and 6-bromo-tryptophan. *GS*-C, heterologous expression of *GS*-C; *GS*-D, heterologous expression of *GS*-D; *GS*-C + D, heterologous expression of *GS*-C and *GS*-D; *GS*-O, heterologous expression of *GS*-O; *GS*-O + E, heterologous expression of *GS*-O and *GS*-E; *GS-B + C + D + A*, and heterologous expression of *GS*-B, *GS*-C, *GS*-D, and *GS*-A. **(B)** Extracted ion chromatograms for the standards (the lower) of indigo (m/z = 261), indirubin (m/z = 261), Tyrian purple (m/z = 421), and colonies were collected from *E. coli* strains expressing different *GS* proteins on the LB agar plates supplemented with tryptophan and 6-bromo-tryptophan.

was observed when feeding 1 mM Trp and 1 mM 6-Br-Trp simultaneously (Figure 5A). LC-MS identified the purple pigment as Tyrian purple, and the common indole transformation products, indigo and indirubin were not detected (Figure 5B). These *in vivo* and *in vitro* reaction results showed that the genes *GS*-A and *GS*-B may relate to the indole degradation, and the four *GS* proteins (*GS*-A, *GS*-B, *GS*-C, and *GS*-D) gene cluster may contribute to the selective Tyrian purple production. Based on the results presented before, *GS*-C was the oxygenase for Tyrian purple biosynthesis without observable selectivity toward indole or 6-Br-indole, and *GS*-D was a cofactor-generating enzyme of *GS*-C. The indole biodegradation pathway containing the four *GS* proteins (*GS*-A, *GS*-B, *GS*-C, and *GS*-D) may contribute to the selective Tyrian purple production.

3.5 Indole biodegradation pathway

To further investigate the molecular mechanisms underpinning this selective Tyrian purple-producing performance, we identified the metabolites involved in the indole biodegradation pathway. A qualitative analysis of the intermediates from indole-degraded extracts was achieved by using liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry (LC-ESI-QTOF-MS/MS) analysis. Several indole degradation bacteria have been isolated and characterized for aerobic biodegradation of indole (Lin et al., 2012; Qu et al., 2015a; Sadauskas et al., 2017; Qu et al., 2017; Ma et al., 2018; Zhang et al., 2018; Ma et al., 2019a). Three major pathways for indole mineralization have been proposed, and these pathways are the catechol pathway, gentisate pathway, and anthranilate pathway (Arora et al., 2015). The proposed compounds related to the indole degradation were tentatively identified from their m/z value and MS spectra in both negative and positive ionization modes ($[M-H]^-/[M+H]^+$) using Agilent LC-MS Qualitative Software and Personal Compound Database and Library (PCDL). Compounds with mass error < ± 5 ppm and PCDL library score more than 80 were selected for further MS/MS identification and m/z characterization purposes (Zhong et al., 2020). As shown in Table 4, the metabolite of isatin was identified according to the comparison with the standard: isatin (Rt = 5.242 min $[M-H]^-$ m/z = 146.0244). Additionally, we identified N-formylanthranilic acid (Rt = 3.599 min $[M-H]^-$ m/z = 164.0355). 2,3-Dihydroxyindole (Rt = 2.213 min $[M+H]^+$ *m*/z = 150.0555), anthranilic acid (Rt = 3.152 min $[M-H]^-$ m/z = 136.0406), salicylic acid (Rt = 4.623 min $[M-H]^-$ m/z = 137.0247), and catechol (Rt = 5.049 min $[M-H]^-$ m/z = 109.0295). Moreover, we observed little response in indoxyl (Rt = 5.389 min; m/z = 133.0529 $[M+Na]^+$ m/z = 156.0431), and indoxyl is a transient product, which is difficult to identify, so further investigation will be needed to identify its chemical structures.

The identified degradation products after indole was completely degraded are consistent with the reported catechol pathway (Arora et al., 2015), so we proposed the indole biodegradation pathway in *Providencia rettgeri* GS-2 is the catechol pathway (Figure 6). Degradation starts with indole oxidation by *GS*-CD at the C-2 and C-3 positions, forming indoxyl. This compound is known to be rather unstable and therefore was not detected. Indoxyl is prone to auto-oxidation and forms an insoluble indigo pigment. However, the auto-oxidation could be prevented by *GS*-B, the hypothetical short-chain dehydrogenase. *GS*-B performs oxidation at the C-2 position to obtain

10.3389/fbioe.2022.1109929

TABLE 4 Results of all product ion scan analyses of biotransformation products related to indole degradation by LC-ESI-QTOF-MS/MS.

Proposed products	Formula	RT/ min	Ionization (ESI^+/ESI^-)	Molecular weight	Theoretical (m/z)	Observed (m/z)	Score	Mass error (ppm)	MS/MS product ions
Indoxyl	C_8H_7NO	1.130		133.0522		156.0413	75.07	-4.77	156.0431 and 56.0497
2,3-Dihydroxyindole	$C_8H_7NO_2$	2.123	$[M + H]^+$	149.0471	150.0549	150.0555	97.7	3.64	132.0447, 122.0626, 104.0498, 94.0640, 77.0387, and 56.0487
Isatin	$C_8H_5NO_2$	4.047	$[M - H]^-$	147.031	146.0236	146.0257	83.14	7.54	118.0301 and 41.9986
N-formylanthranilic acid	$C_8H_7NO_3$	3.599	$[M - H]^-$	165.0420	164.0342	164.0355	99.49	1.43	92.0506 and 120.0455
Anthranilic acid	$C_7H_7NO_2$	3.599	$[M - H]^-$	137.0471	136.0393	136.0406	86.55	1.67	92.0502 and 42.0357
Salicylic acid	$C_7H_6O_3$	4.623	$[M - H]^-$	138.0311	137.0233	137.0247	98.65	1.84	93.0354, 65.0393, and 50.0035
Catechol	$C_6H_6O_2$	5.049		110.0362	109.0284	109.0295	85.98	0.19	

* Proposed products were detected in both negative $[M-H]^-$ and positive $[M + H]^+$ mode of ionization. RT = stands for "retention time". Theoretical (m/z) stands for the theoretical m/z under certain mode of ionization. Observed (m/z) stands for the detected m/z. Mass error stands for the difference between theoretical m/z and observed m/z in ppm.

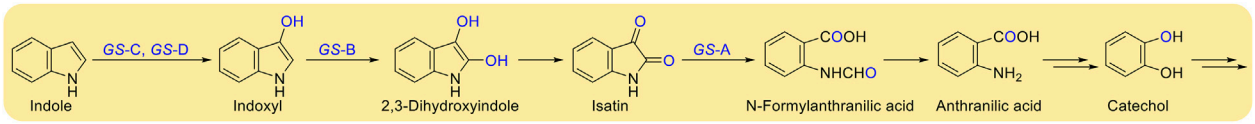


FIGURE 6
Proposed indole biodegradation pathway in *Providencia rettgeri* GS-2.

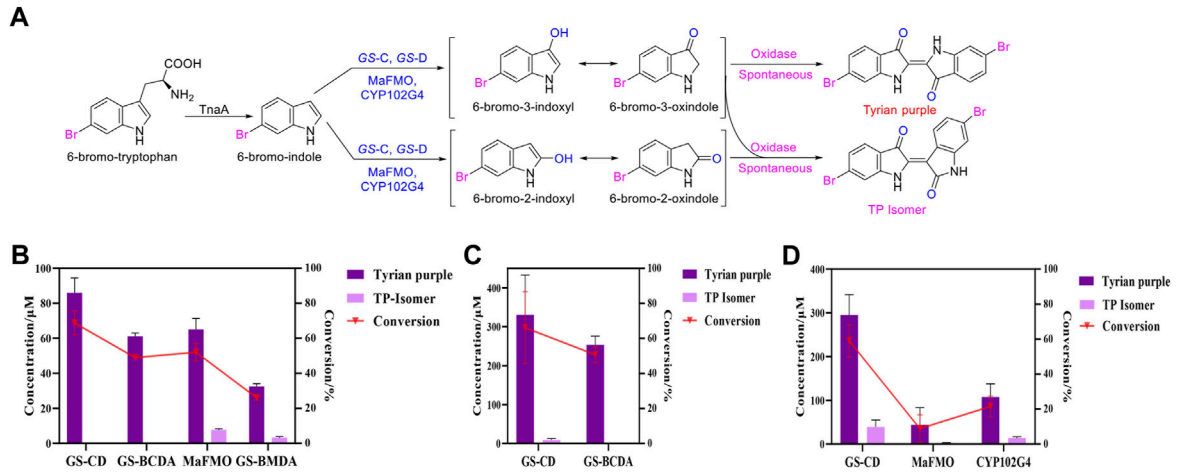


FIGURE 7
Production of Tyrian purple and TP isomer from 6-Br-tryptophan or 6-Br-indole using *E. coli*, expressing flavin-containing monooxygenase from *Providencia rettgeri* GS-2, *Methylophaga aminisulfidivorans* (MaFMO), or cytochrome P450 enzyme CYP102G4. **(A)** Scheme of synthesis of Tyrian purple and TP isomer from 6-bromo-tryptophan. **(B)** Production and conversion of Tyrian purple by feeding 250 μM 6-Br-indole substrate *in vivo*. **(C)** Production and conversion of Tyrian purple by feeding 1,000 μM 6-Br-indole substrate in whole-cell reaction. **(D)** Production and conversion of Tyrian purple by feeding 1,000 μM 6-Br-tryptophan substrate *in vivo*.

2, 3-dihydroxyindole. 2, 3-Dihydroxyindole is spontaneous to forming a stable isatin intermediate. N-formylanthranilic acid, the yellow product, which is catalyzed by *GS*-A, is then further degraded to anthranilic acid. Anthranilic acid was ultimately generated as the typical downstream product, which would be further degraded *via* salicylic acid and catechol. As it stands, the proposed indole

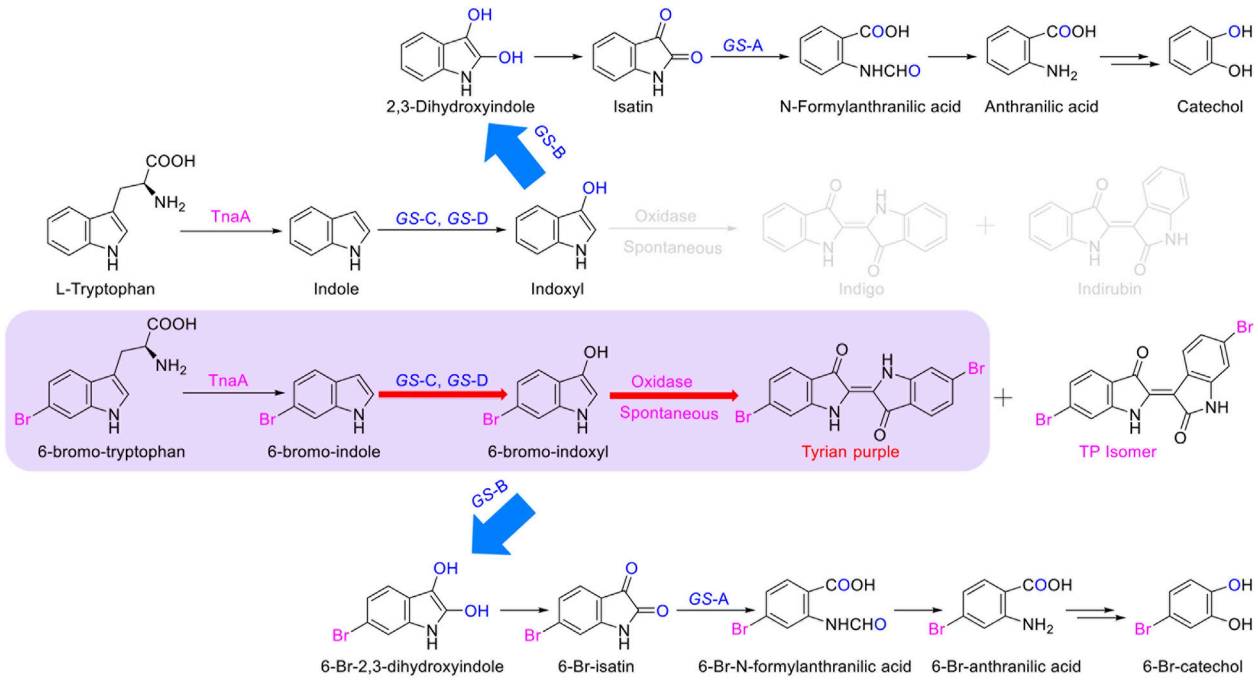


FIGURE 8
Proposed metabolic pathways contribute to the selectivity Tyrian purple-producing in *E. coli*.

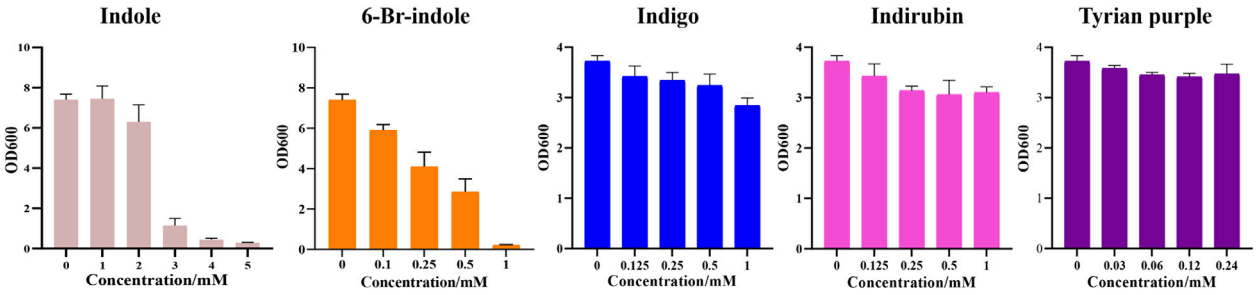


FIGURE 9
Effect of indole, 6-Br-indole, indigo, indirubin, and Tyrian purple on the growth of *E. coli*.

biodegradation pathway in *Providencia rettgeri* GS-2 is the catechol pathway that requires additional experiments to prove the role of putative catalytic by *GS* proteins (*GS*-A, *GS*-B, *GS*-C, and *GS*-D), and these experiments are underway.

3.6 Production of Tyrian purple in *E. coli*

Tyrian purple is an expensive dye and has a promising application in modern times. MaFMO has been reported as an effective oxygenase for Tyrian purple biosynthesis by the oxidation of 6-Br-indole (Lee et al., 2021; Schnepel et al., 2021). The activity of *GS*-C and the indole degradation gene cluster (four *GS* proteins *GS*-A, *GS*-B, *GS*-C, and *GS*-D) contribute to the production of Tyrian purple. We synthesized MaFMO and expressed in *E. coli* BL21 (DE3) with or without the genes *GS*-A, *GS*-B, and *GS*-D in the indole degradation cluster (designed as MaFMO or *GS*-BMDA). We further compared the Tyrian purple and TP isomer production by feeding 250 μM 6-Br-indole substrate in M9 minimal medium. LC-MS-identified *GS*-CD (*GS*-C and *GS*-D) have the highest Tyrian purple production with 85.9 μM and 68.7% conversion ratio *in vivo* (Figure 7B). With *GS*-BA (*GS*-A and *GS*-B), the production of Tyrian purple decreased to 61.2 μM, and the conversion ratio was 49%. Meanwhile, no TP isomer has been detected in the products. The production and conversion ratio of Tyrian purple by MaFMO was 65 μM and 52%, respectively. With *GS*-BA (*GS*-A and *GS*-B), the production of Tyrian purple decreased to 32.5 μM, and the conversion ratio was 26%. Meanwhile, a part of TP isomer has been detected in the products of MaFMO with or without *GS*-BA. These results confirm that the Tyrian purple-producing enzymes *GS*-CD have better activity and selectivity (with *GS*-BA) in the production of Tyrian purple in *E. coli* than MaFMO. When feeding the 1,000 μM 6-Br-indole substrate in a whole-cell reaction (Figure 7C), the *GS*-CD has the highest Tyrian

TABLE 5 Steady-state kinetic parameters for NADH oxidation activity on various substrates by *GS*-CD and MaFMO.

Protein	Substrate	K_m/μm	K_{cat}/min	K_{cat}/K_m
GS-CD	Indole	217.4	0.05796	0.00027
	6-Br-indole	405.8	0.1264	0.00031
MaFMO	Indole	242.2	0.01475	0.00006
	6-Br-indole	303.65	0.03129	0.00010

purple production with 330.8 μM and 66.2% conversion ratio in the M9 medium. With *GS-BA*, the production of Tyrian purple decreased by 24% compared with *GS-CD* in the M9 medium. Therefore, by taking advantage of the indole degradation gene cluster in GS-2, we showed that up to 253.4 μM Tyrian purple could be selectively produced from 1,000 μM 6-Br-indole.

As shown in Supplementary Figure 1B, heterologous expression of *GS*-B, *GS*-C, *GS*-D, and *GS*-A from GS-2 in *E. coli* have better production of Tyrian purple than GS-2 when 6-Br-Trp was used as a substrate on a solid plate. To further investigate the biosynthesis of Tyrian purple from 6-Br-Trp in *E. coli*, cytochrome P450 enzyme CYP102G4 (Namgung et al., 2019) was synthesized (CYP102G4 was reported to effectively synthesize indigoid dyes by biotransform indole derivatives). To produce Tyrian purple from 6-Br-Trp, each monooxygenase was co-expressed with TnaA, resulting in three strains, ΔtnaA TnaA + *GS*-CD (designed as *GS*-CD), ΔtnaA TnaA + MaFMO (designed as MaFMO), and ΔtnaA TnaA + CYP102G4 (designed as CYP102G4), and their expression conditions were optimized. When feeding 1,000 μM 6-Br-Trp substrate *in vivo*, compared with MaFMO and CYP102G4, the *GS*-CD has the highest Tyrian purple production with 295.4 μM and 59.1% conversion ratio *in vivo* (Figure 7D). The production and conversion ratio of Tyrian purple by MaFMO were 44.3 μM and 8.9%, respectively. The production and conversion ratio of Tyrian purple by CYP102G4 were 107.5 μM and 21.5%, respectively. Therefore, the flavin-containing monooxygenase *GS*-CD from *Providencia rettgeri* GS-2 is an effective Tyrian purple-producing enzyme, and the indole degradation gene cluster *GS*-BCDA can be used for the selective production of Tyrian purple in *E. coli*.

4 Discussion

In this study, we identified that the indole biodegradation gene cluster from *Providencia rettgeri* GS-2 could contribute to the selective Tyrian purple production. The molecular mechanisms underpinning this unique performance warrant further investigation. We proposed the indole biodegradation pathway in *Providencia rettgeri* GS-2, which is the catechol pathway according to the degradation products of indole. Notably, the indole biodegradation gene cluster can transform and biodegrade both indole and 6-Br-indole practically. On the one hand, the indole oxygenase *GS*-CD can utilize both indole and 6-Br-indole to form insoluble indigo or Tyrian purple pigment. On the other hand, the oxidations formed by *GS*-CD can be further degraded by *GS*-B and *GS*-A (Figure 8). Based on the *in vivo* and *in vitro* results presented before, *GS*-C and its cofactor-generating enzyme *GS*-D are responsible for the Tyrian purple biosynthesis without observable selectivity toward indole or 6-Br-indole. The indole biodegradation genes *GS*-A and *GS*-B may contribute to the selective Tyrian purple production. Even though we have purified *GS*-A and *GS*-B, it is difficult to obtain the standards of 2,3-dihydroxyindole and N-formylanthranilic acid,

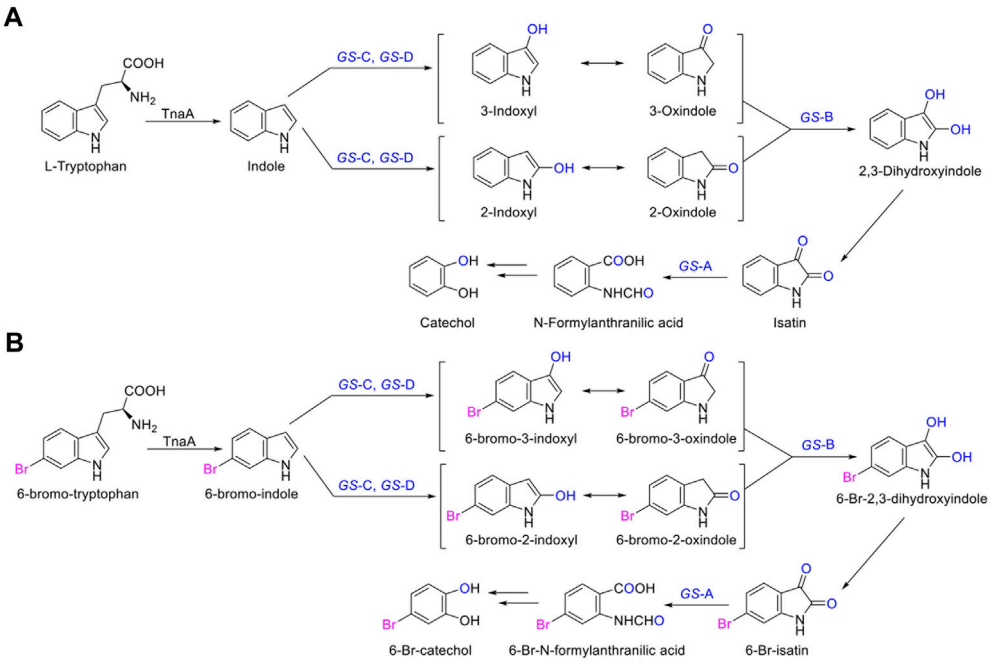


FIGURE 10
(A,B) Scheme of the indole biodegradation gene clusters assisted selective synthesis of Tyrian purple.

the further investigation of the molecular mechanisms of the selective Tyrian purple production would jog along.

Furthermore, during our substrate supplement experiments, we found that 6-Br-indole has a significant influence on the growth of *E. coli*. By comparing the effect of indole, 6-Br-indole, indigo, indirubin, and Tyrian purple on the growth of *E. coli*, we found 6-Br-indole has a significant influence on the growth of *E. coli* even at the concentration of 100 μM (Figure 9). The toxicity of indole is above 2 mM. Indigo, indirubin, and Tyrian purple have little effect on the growth of *E. coli*. Thus, the toxicity of 6-Br-indole may drive the faster transformation of 6-Br-indole to non-toxic Tyrian purple. Meanwhile, the catalytic efficiency of *GS*-CD (k_{cat}) on 6-Br-indole is faster than that of indole (about two-folds) (Table 5). The transformation of 6-Br-indole to Tyrian purple probably is dominant, and the biodegradation of 6-Br-indole is minor. As the toxicity of indole is much smaller than 6-Br-indole, and the catalytic efficiency of *GS*-CD (k_{cat}) on indole is slow, the strain may maintain a balance between indole transformation and degradation, that is, indole was catalyzed by the *GS*-CD to indoxyl, which was rapidly degraded by *GS*-B and *GS*-A, so indigo/indirubin cannot be synthesized. Therefore, we identified Tyrian purple, the common indole transformation product, and indigo and indirubin were not detected when the indole biodegradation gene cluster reaction *in vivo* or *in vitro* took place (Figures 4C, 5).

In the biosynthesis of Tyrian purple, the formation of TP isomer, which is a stereoisomer of Tyrian purple, is a major side reaction. In addition, the indole oxygenase *GS*-C oxidized the C-2 and C-3 positions of 6-Br-indole, forming the C-2 and C-3 6-Br-indoxyl, which can be an auto-oxidation to Tyrian purple and TP isomer simultaneously (Figure 7). As C2-specific hydroxylation of indole/6-Br-indole leads to the formation of isatin/6-Br-isatin, which further reacts with indoxyl/6-Br-indoxyl and generates the asymmetrical form of indirubin/TP isomer (Cooksey, 2001). In the proposed indole biodegradation pathway, indole biodegradation genes *GS*-B and *GS*-A can degrade C-2 and C-3 oxidations of 6-Br-indole to 6-Br-isatin simultaneously, and further be degraded (Figure 10), but the amount of C-3 oxidation is far more than C-2's (Figures 4C, 7C). So the C-2 oxidation-related transformation product TP isomer cannot be detected. Thus, the indole biodegradation gene cluster biodegrades isatin/6-Br-isatin contribute to the selective Tyrian purple production.

We also confirmed that the Tyrian purple-producing enzymes *GS*-CD have better activity in the production of Tyrian purple in *E. coli* than the most reported monooxygenase MaFMO. The steady-state kinetic parameters for NADH oxidation activity on various substrates by *GS*-CD and MaFMO proteins were determined (Table 5). The catalytic efficiency of *GS*-CD (k_{cat}/K_m) on indole/6-Br-indole was higher than that of MaFMO. Those results further prove that *GS*-CD have better activity in the production of Tyrian purple in *E. coli* than MaFMO. The effective enzymes *GS*-CD that are found would lay a solid foundation for the promising biosynthesis of Tyrian purple. Incorporating the indole biodegradation gene cluster for the production of Tyrian purple would not only allow obtaining a purer product but also simplify the downstream steps of biosynthesis.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI SUB12350614.

Author contributions

FL, XZ, and XL conceived the study. FL constructed the plasmids and *E. coli* strains and performed microbiological manipulations and extractions. FL and BZ performed mass spectrometry. FL and BR performed enzyme kinetic study and *in vitro* studies. FL, XZ, and HD participated in data collection and manuscript preparation. XZ, HD, and XL, reviewed the manuscript. All authors played an important role in this work and approved it to be published.

Funding

This study was financially supported by the Shenzhen Science and Technology Program (ZDSYS20210623091810032) and Shenzhen Institute of Synthetic Biology Scientific Research Program (ZTXM20203001). The authors also acknowledge the fruitful discussions from prof. Buhe Nashun, Dr. Que Chen, and Dr. Zhilai Hong.

Conflict of interest

XL has a financial interest in Demetrix and Synceres.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.1109929/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Growth curve and the yields of Tyrian purple-producing strains. **(A)** Growth curve of Tyrian purple-producing strains. **(B)** Yields of Tyrian purple produced by Tyrian purple-producing strains when feeding different substrates on a solid plate. **GS-2**, in the bacterial strain GS-2. ***E. coli***, heterologous expression of *GS*-B, *GS*-C, *GS*-D, and *GS*-A from GS-2 in *E. coli*.

SUPPLEMENTARY FIGURE S2
Flavin reductase activity of the purified *GS*-D protein.

References

Alemayehu, D., Gordon, L.M., O'mahony, M.M., O'leary, N.D., and Dobson, A.D. (2004). Cloning and functional analysis by gene disruption of a novel gene involved in indigo production and fluoranthene metabolism in Pseudomonas alcaligenes PA-10. *FEMS Microbiol Lett* 239, 285–293. doi:10.1016/j.femsle.2004.08.046

Ameria, S.P., Jung, H.S., Kim, H.S., Han, S.S., Kim, H.S., and Lee, J.H. (2015). Characterization of a flavin-containing monooxygenase from Corynebacterium glutamicum and its application to production of indigo and indirubin. *Biotechnol Lett* 37, 1637–1644. doi:10.1007/s10529-015-1824-2

Doukyu, N, and Aono, R. (1997). Biodegradation of indole at high concentration by persolvent fermentation with Pseudomonas sp. ST-200. *Extremophiles* 1, 100–105. doi:10.1007/s007920050021

Arora, P.K., Sharma, A., and Bae, H. (2015). Microbial Degradation of Indole and Its Derivatives. *J. Chem.* 2015, 1–13. doi:10.1155/2015/129159

Bhushan, B., Samanta, S. K., and Jain, R. (2000). Indigo production by naphthalene-degrading bacteria. *Lett. Appl. Microbiol.* 31, 5–9. doi:10.1046/j.1472-765x.2000.00754.x

O'hara, C. M., Brenner, F. W., and Miller, J. M. (2000). Classification, Identification, and Clinical Significance of Proteus, Providencia, and Morganella. *Clin. Microbiol. Rev.* 13 (4), 534–46. doi:10.1128/CMR.13.4.534

Choi, K.-Y. (2020). A review of recent progress in the synthesis of bio-indigoids and their biologically assisted end-use applications. *Dyes. Pigm* 181, 108570. doi:10.1016/j.dyepig.2020.108570

Cooksey, C.J. (2001). Tyrian Purple 6, 6'-Dibromoindigo and Related Compounds. *Molecules* 6, 736–769. doi:10.3390/60900736

Dai, C., Ma, Q., Li, Y., Zhou, D., Yang, B., and Qu, Y. (2019). Application of an efficient indole oxygenase system from Cupriavidus sp. SHE for indigo production. *Bioprocess. Biosyst. Eng.* 42, 1963–1971. doi:10.1007/s00449-019-02189-4

Doukyu, N., Toyoda, K., and Aono, R. (2003). Indigo production by Escherichia coli carrying the phenol hydroxylase gene from Acinetobacter sp strain ST-550 in a water-organic solvent two-phase system. *Appl. Microbiol. Biotechnol.* 60, 720–725. doi:10.1007/s00253-002-1187-1

Drewlo, S., Bramer, C.O., Madkour, M., Mayer, F., and Steinbuchel, A. (2001). Cloning and expression of a Ralstonia eutropha HF39 gene mediating indigo formation in Escherichia coli. *Appl. Environ. Microbiol.* 67, 1964–1969. doi:10.1128/AEM.67.4.1964-1969.2001

Fabara, A.N., and Fraaije, M.W. (2020). An overview of microbial indigo-forming enzymes. *Appl. Microbiol. Biotechnol.* 104, 925–933. doi:10.1007/s00253-019-10292-5

Fukuoka, K., Tanaka, K., Ozeki, Y., and Kanaly, R.A. (2015). Biotransformation of indole by Cupriavidus sp. strain KK10 proceeds through N-heterocyclic- and carbocyclic-aromatic ring cleavage and production of indigoids. *Int. Biodeterior. Biodegradation.* 97, 13–24. doi:10.1016/j.ibiod.2014.11.007

Galac, M., and Lazzaro, B. (2012). Comparative genomics of bacteria in the genus Providencia isolated from wild Drosophila melanogaster. *BMC Genomics* 13, 612. doi:10.1186/1471-2164-13-612

Głowacki, E.D., Voss, G., Leonat, L., Irimia-Vladu, M., Bauer, S., and Sariciftci, N.S. (2012). Indigo and Tyrian Purple - From Ancient Natural Dyes to Modern Organic Semiconductors. *Isr. J. Chem.* 52, 540–551. doi:10.1002/ijch.201100130

Głowacki, E.D., Voss, G., and Sariciftci, N.S. (2013). 25th Anniversary Article: Progress in Chemistry and Applications of Functional Indigos for Organic Electronics. *Adv. Mater.* 25, 6783–6800. doi:10.1002/adma.201302652

Guo, C., Quinn, J., Sun, B., and Li, Y. (2015). An indigo-based polymer bearing thermocleavable side chains for n-type organic thin film transistors. *J. Mater. Chem. C.* 3, 5226–5232. doi:10.1039/c5tc00512d

Heine, T., van Berkel, W. J. H., Gassner, G., van Pée, K. H., and Tischler, D. (2018). Two-Component FAD-Dependent Monooxygenases: Current Knowledge and Biotechnological Opportunities. *Biology* 7, 42. doi:10.3390/biology7030042

Heine, T., Grossmann, C., Hofmann, S., and Tischler, D. (2019). Indigoid dyes by group E monooxygenases: mechanism and biocatalysis. *Biol. Chem.* 400, 939–950. doi:10.1515/hsz-2019-0109

Kim, J., Lee, K., Kim, Y., Kim, C., and Lee, K. (2003). Production of dyestuffs from indole derivatives by naphthalene dioxygenase and toluene dioxygenase. *Lett Appl Microbiol* 36, 343–348. doi:10.1046/j.1472-765x.2003.01279.x

Kim, H., Kim, G., Song, I., Lee, J., Abdullah, H., Yang, C., and Oh, J.H. (2018). Ambipolar organic phototransistors based on 6, 6'-dibromoindigo. *RSC Adv* 8, 14747–14752. doi:10.1039/c8ra02346h

Kim, J., Lee, J., Lee, P.-G., Kim, E.-J., Kroutil, W., and Kim, B.-G. (2019). Elucidating Cysteine-Assisted Synthesis of Indirubin by a Flavin-Containing Monooxygenase. *ACS Catal* 9, 9539–9544. doi:10.1021/acscatal.9b02613

Kim, M., Lee, J.H., Kim, E., Choi, H., Kim, Y., and Lee, J. (2016). Isolation of Indole Utilizing Bacteria Arthrobacter sp. and Alcaligenes sp. From Livestock Waste. *Indian J. Microbiol.* 56, 158–166. doi:10.1007/s12088-016-0570-z

Kwon, N.R., Chae, J.C., Choi, K.Y., Yoo, M., Zylstra, G.J., Kim, Y.M., Kang, B.S., and Kim, E. (2008). Identification of functionally important amino acids in a novel indigo-producing oxygenase from Rhodococcus sp. strain T104. *Appl. Microbiol. Biotechnol.* 79, 417–422. doi:10.1007/s00253-008-1445-y

Lee, J., Kim, J., Song, J.E., Song, W.S., Kim, E.J., Kim, Y.G., Jeong, H.J., Kim, H.R., Choi, K.Y., and Kim, B.G. (2021). Production of Tyrian purple indigoid dye from tryptophan in Escherichia coli. *Nat. Chem. Biol.* 17 (1), 104–112. doi:10.1038/s41589-020-00684-4

Lee, J.H., and Lee, J. (2010). Indole as an intercellular signal in microbial communities. *FEMS Microbiol. Rev.* 34, 426–444. doi:10.1111/j.1574-6976.2009.00204.x

Li, X., Zhang, B., Hu, Y., and Zhao, Y. (2021). New Insights Into Gut-Bacteria-Derived Indole and Its Derivatives in Intestinal and Liver Diseases. *Front. Pharmacol.* 12, 769501. doi:10.3389/fphar.2021.769501

Lin, G.H., Chen, H.P., Huang, J.H., Liu, T.T., Lin, T.K., Wang, S.J., Tseng, C.H., and Shu, H.Y. (2012). Identification and characterization of an indigo-producing oxygenase involved in indole 3-acetic acid utilization by Acinetobacter baumannii. *Antonie. Van. Leeuwenhoek* 101 (4), 881–890. doi:10.1007/s10482-012-9704-4

Lin, G.H., Chen, H.P., and Shu, H.Y. (2015). Detoxification of Indole by an Indole-Induced Flavoprotein Oxygenase from Acinetobacter baumannii. *PLoS One* 10 (9), e0138798. doi:10.1371/journal.pone.0138798

Lu, Y., and Mei, L. (2007). Co-expression of P450 BM3 and glucose dehydrogenase by recombinant Escherichia coli and its application in an NADPH-dependent indigo production system. *J. Ind. Microbiol. Biotechnol.* 34, 247–253. doi:10.1007/s10295-006-0193-1

Ma, Q., Liu, Z., Yang, B., Dai, C., and Qu, Y. (2019a). Characterization and functional gene analysis of a newly isolated indole-degrading bacterium Burkholderia sp. IDO3. *J. Hazard. Mater.* 367367, 144144–151151. doi:10.1016/j.jhazmat.2018.12.068

Ma, Q., Yang, B., Qu, H., Gao, Z., Qu, Y., and Sun, Y. (2019b). Identification and functional study of an iif2 gene cluster for indole degradation in Burkholderia sp. IDO3. *Int. Biodeterior. Biodegradation.* 142, 36–42. doi:10.1016/j.ibiod.2019.04.011

Ma, Q., Zhang, X., and Qu, Y. (2018). Biodegradation and Biotransformation of Indole: Advances and Perspectives. *Front. Microbiol.* 9, 2625. doi:10.3389/fmicb.2018.02625

Mcgovern, P., and Michel, R. (1990). Royal purple dye: the chemical reconstruction of the ancient Mediterranean industry. *Acc. Chem. Res.* 23, 152–158. doi:10.1021/ar00173a006

Mendoza-Avila, J., Chauhan, K., and Vazquez-Duhalt, R. (2020). Enzymatic synthesis of indigo-derivative industrial Dyes. *Pigm* 178, 108384. doi:10.1016/j.dyepig.2020.108384

Namgung, S., Park, H.A., Kim, J., Lee, P.-G., Kim, B.-G., Yang, Y.-H., and Choi, K.-Y. (2019). Ecofriendly one-pot biosynthesis of indigo derivative dyes using CYP102G4 and PrnA halogenase. *Dyes Pigm* 162, 80–88. doi:10.1016/j.dyepig.2018.10.009

Ngangbam, A.K., Waters, D.L.E., Whalan, S., Baten, A., and Benkendorff, K. (2015). Indole-Producing Bacteria from the Biosynthetic Organs of a Muricid Mollusc Could Contribute to Tyrian Purple Production. *J. Shellfish Res.* 34 (2), 443–454. doi:10.2983/035.034.0228

O'Connor, K. E., Dobson, A. D., and Hartmans, S. (1997). Indigo Formation by Microorganisms Expressing Styrene Monooxygenase Activity. *Appl. Environ. Microbiol.* 63 (11), 4287–4291. doi:10.1128/aem.63.11.4287-4291.1997

Qu, Y., Ma, Q., Liu, Z., Wang, W., Tang, H., Zhou, J., and Xu, P. (2017). Unveiling the biotransformation mechanism of indole in a Cupriavidus sp. strain. *Mol. Microbiol.* 106 (6), 905–918. doi:10.1111/mmi.13852

Qu, Y., Shen, E., Ma, Q., Zhang, Z., Liu, Z., Shen, W., Wang, J., Li, D., Li, H., and Zhou, J. (2015a). Biodegradation of indole by a newly isolated Cupriavidus sp. SHE. *J. Environ. Sci (China).* 34, 126–132. doi:10.1016/j.jes.2015.01.023

Qu, Y., Shi, S., Zhou, H., Ma, Q., Li, X., Zhang, X., and Zhou, J. (2012). Characterization of a Novel Phenol Hydroxylase in Indoles Biotranformation from a Strain Arthrobacter sp. W1. *PLoS One* 7, e44313. doi:10.1371/journal.pone.0044313

Qu, Y., Zhang, Z., Ma, Q., Shen, E., Shen, W., Wang, J., Cong, L., Li, D., Liu, Z., Li, H., and Zhou, J. (2015b). Biotransformation of Indole and Its Derivatives by a Newly Isolated Enterobacter sp. M9Z. *Appl. Biochem. Biotechnol.* 175, 3468–3478. doi:10.1007/s12010-015-1518-1

Sadauskas, M, Vaitekūnas, J, Gasparavičiūtė, R, and Meškys, R (2017). Indole Biodegradation in Acinetobacter sp. Strain O153: Genetic and Biochemical Characterization. *Appl. Environ. Microbiol.* 83 (19), e01453–17. doi:10.1128/AEM.01453-17

Schnepel, C., Dodero, V.I., and Sewald, N. (2021). Novel Arylindigoids by Late-Stage Derivatization of Biocatalytically Synthesized Dibromoindigo. *Chem. Eur. J.* 27, 5404–5411. doi:10.1002/chem.202005191

Tomberlin, J.K., Crippen, T.L., Wu, G., Griffin, A.S., Wood, T.K., and Kilner, R.M. (2017). Indole: An evolutionarily conserved influencer of behavior across kingdoms. *Bioessays* 39 (2), 1600203. doi:10.1002/bies.201600203

Wolk, J.L., and Frimer, A.A. (2010). Preparation of Tyrian purple (6, 6'-dibromoindigo): past and present. *Molecules* 15, 5473–5508. doi:10.3390/molecules15085473

Yin, B., Gu, J.-D., and Wan, N. (2005). Degradation of indole by enrichment culture and Pseudomonas aeruginosa Gs isolated from mangrove sediment. *Int. Biodeterior. Biodegradation.* 56, 243–248. doi:10.1016/j.ibiod.2005.10.001

Zhang, X., Jing, J., Zhang, L., Song, Z., Zhou, H., Wu, M., Qu, Y., and Liu, L. (2018). Biodegradation characteristics and genomic functional analysis of indole-degrading bacterial strain *Acinetobacter* sp. JW. *J. Chem. Technol. Biotechnol.* 94, 1114–1122. doi:10.1002/jctb.5858

Zhang, X., Qu, Y., Ma, Q., Zhou, H., Li, X., Kong, C., and Zhou, J. (2013). Cloning and expression of naphthalene dioxygenase genes from Comamonas sp. MQ for indigoids production. *Process Biochem* 48, 581–587. doi:10.1016/j.procbio.2013.02.008

Zhong, B., Robinson, N.A., Warner, R.D., Barrow, C.J., Dunshea, F.R., and Suleria, H.a.R. (2020). LC-ESI-QTOF-MS/MS Characterization of Seaweed Phenolics and Their Antioxidant Potential. *Mar. Drugs.* 18 (6), 331. doi:10.3390/md18060331

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 26 January 2023
DOI 10.3389/fbioe.2023.1131875

OPEN ACCESS

EDITED BY
Zhenlin Han,
University of Hawaii at Manoa,
United States

REVIEWED BY
Xiumei Li,
Feed Research Institute, Chinese Academy of Agricultural Sciences, China
Xianpu Ni,
Shenyang Pharmaceutical University, China

*CORRESPONDENCE
Wenqin Bai,
baiwq@tib.cas.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and Biotechnology

RECEIVED 26 December 2022
ACCEPTED 19 January 2023
PUBLISHED 26 January 2023

CITATION
Chen S, Zheng H, Gao J, Song H and Bai W (2023), High-level production of pullulan and its biosynthesis regulation in *Aureobasidium pullulans* BL06.
Front. Bioeng. Biotechnol. 11:1131875.
doi: 10.3389/fbioe.2023.1131875

High-level production of pullulan and its biosynthesis regulation in *Aureobasidium pullulans* BL06

Shuyu Chen[1†], Hongchen Zheng[2,3,4,5†], Jiaqi Gao[2,3,4], Hui Song[2,3,4] and Wenqin Bai[2,3,4,5*]

[1]Colleg of Biotechnology, Tianjin University of Science Technology, Tianjin, China, [2]CAS Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China, [3]National Center of Technology Innovation for Synthetic Biology, Tianjin, China, [4]Industrial Enzymes National Engineering Research Center, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China, [5]Key Laboratory of Engineering Biology for Low-carbon Manufacturing, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China

Pullulan has many potential applications in the food, pharmaceutical, cosmetic and environmental industries. However, the yield and molecular properties of pullulan produced by various strains still need to be promoted to fit the application needs. A novel yeast-like strain *Aureobasidium pullulans* BL06 producing high molecular weight (Mw) pullulan (3.3×10^6 Da) was isolated and identified in this study. The remarkable Mw of pullulan produced by *A. pullulans* BL06 was the highest level ever reported thus far. To further regulate the biosynthesis of pullulan in *A. pullulans* BL06, three gene knockout strains *A. pullulans* BL06 ΔPMAs, *A. pullulans* BL06 Δmel, and *A. pullulans* BL06 ΔPMAsΔmel, were constructed. The results showed that *A. pullulans* BL06 ΔPMAs could produce 140.2 g/L of moderate Mw (1.3×10^5 Da) pullulan after 120 h of fermentation. The highest yield level of pullulan to date could vastly reduce its production cost and expand its application scope and potential. The application experiments in food preservation showed that the moderate-Mw pullulan obtained in this work could reduce the weight loss of celery cabbages and mangos by 12.5% and 22%, respectively. Thus, the novel strains *A. pullulans* BL06 and *A. pullulans* BL06 ΔPMAs possessed unlimited development prospects in pullulan production at various Mw ranges and pullulan applications in multiple fields.

KEYWORDS

pullulan, aureobasidium pullulans, biosynthesis, high-level production, high molecular weight, moderate molecular weight, food preservation

1 Introduction

Pullulan, composed of repeating maltotriose units connected by α-(1, 6) glycosidic bonds, is a linear extracellular polysaccharide (EPS) that is usually produced by the yeast-like fungus *Aureobasidium pullulans* (*A. pullulans*) (Singh et al., 2021; Wei et al., 2021). It exhibits a wide range of applications in biomedicine, environmental engineering, and food engineering because it is a renewable, toxin-free, non-immunogenic and non-carcinogenic natural polymer (Singh et al., 2019; Nishimura et al., 2020; An et al., 2021; Feng et al., 2022). More specifically, the unique linkage pattern ofα-D-glucan endows pullulan with excellent solubility, film-forming ability, oxygen barrier ability, and structural elasticity to meet the requirements for application in drug delivery, gene targeting, tissue engineering, wound healing, and 3D/4D printing (Duan et al., 2020; Ghorbani et al., 2020; Moraes et al., 2020; Shah et al., 2021; Feng et al., 2022). The intramolecular and intermolecular polyhydroxyl interactions of the linear pullulan molecule are key to its superb material properties. High-molecular-weight (High-Mw) pullulan could offer stronger mechanical properties due to its longer polysaccharide

chain and its richer hydroxyl content, which makes it a promising candidate in biomedical materials and shape memory polymers (SMPs) (Jiang et al., 2019; Feng et al., 2022). However, pullulan-based hard capsules were recently manufactured by using commercial pullulan (e.g., Hayashibara Co., Ltd., Japan) with a mean molecular weight (Mw) of 1.0×10^5 or 2.0×10^5 Da (Liu et al., 2018; Ding et al., 2020). Besides, to date, the application of pullulan in various fields has been actually limited due to both the low production yield and low Mw of pullulan (Liu et al., 2018). Owing to the importance of the properties and production cost of pullulan, bioprocessing and strain modification have been widely studied to enhance the Mw and yield of pullulan. Additionally, efficient genome editing techniques which were established in 2019 also provide the assistance to regulate the relevant enzymes and genes of the pullulan biosynthesis pathway in different strains of *Aureobasidium* spp. (Zhang et al., 2019). For example, simultaneous removal of both duplicated AMY1 genes encoding α-amylase and duplicated PKS1 genes responsible for melanin biosynthesis in *A. melanogenum* TN3-1 rendered a mutant AMY-PKS-11 to transform 140.0 g/L glucose to produce 103.50 g/L pigment-free pullulan with a Mw of 3.2×10^5 Da (Chen et al., 2019; Xue et al., 2019; Chen et al., 2020). A triple mutant DT15 grown at the flask level could produce 46.2 g/L of pullulan with a Mw of 3.02×10^6 Da and grown in a 10-L fermentor could yield 58.14 g/L pullulan with the same Mw, while its wild-type strain P16 produced 65.5 ± 3.5 g/L pullulan with a Mw of 0.35×10^6 Da (Liu et al., 2018). Thus, more efforts should be made to enhance pullulan production and improve the chemical properties of pullulan *via* molecular modifications of the producers by using synthetic biology approaches.

Pullulan produced by *A. pullulans* fermentation normally contains byproducts, including melanin, glucan and polymalic acid (Zeng et al., 2020; Chen et al., 2021; Liu et al., 2021). The presence of these impurities largely increases the difficulty of pullulan purification. Due to differences in metabolic pathways and cell morphology, the biosynthetic mechanisms vary by different strains and therefore cause the varied product molecular weight among different strains (Sugumaran and Ponnusami, 2017). In this work, the high yield pullulan producing strain *A. pullulans* BL06 was screened from the environment. Pullulan synthesized by this strain showed a molecular weight of 3.3×10^6 Da and a fermentation yield of 83.4 g/L in a 5 L bioreactor. The Mw of the novel identified pullulan is the highest among those of the pullulan ever reported. The strain BL06 combines the advantages of both high molecular weight and high yield and therefore has a high potential for commercialization. Furthermore, by knocking out the polymalic acid (PMA) synthase gene, we obtained another industrial strain *A. pullulans* BL06 ΔPMAs that are capable of producing high purity and moderate Mw pullulan with a high yield. The fermentation of the modified strain reached 140.2 g/L yield of 1.3×10^5 Da pullulan in a 5 L bioreactor, free of melanin and PMA impurities. The application experiments in food preservation of the moderate Mw pullulan obtained in this work were also performed to evaluate its developing potential in the food industry.

2 Materials and methods

2.1 Strains, plasmids, and chemicals

All of the wild strains producing exopolysaccharides in this study were isolated from the fallen leaves of the park (N39°8′E117°23′) near the Tianjin institute of industrial biotechnology, Chinese academy of sciences in September 2020. All the plasmids used in this work are listed in Supplementary Table S1. *E. coli* DH5α was used to preserve and amplify the recombinant shuttle plasmids. Maltotriose standard, pullulan standard and pullulanase were purchased from Sigma (St. Louis, MO, United States). All other chemicals were of analytical reagent grade purity and obtained from commercial sources.

2.2 Screening of the pullulan-producing strains

According to previous reports, most of the pullulan producers were isolated from various plant leaves and flowers (Ma et al., 2014). Four kinds of fallen leaves from different trees were used as the sources for fungus isolation in this study. The detailed location information was described in Section 2.1. The fungal-carrying sample obtained from the environment was washed and serially diluted with 0.9% sodium chloride solution and then spread on a potato dextrose agar (PDA) plate, which was 100 mL of potato extract containing 12.0 g of sucrose and 2.0 g of agar with a certain amount of bacterial antibiotics (ampicillin 140 μg/mL, chloromycetin 200 μg/mL), followed by incubation at 28°C for 24 h. After incubation and colony formation, single colonies which without pigment generation and showed smooth and moist in the surface were chosen to identify the production ability of EPS. The yeast-like properties of the single colonies were referred to the standard colonial morphology as shown in Figure 1C. The yeast-like fungal single colonies were further aerobically cultivated in YPD medium (consisting of 1.0% yeast extract, 2.0% polypeptone and 2.0% glucose) at 28°C for 24 h to obtain seed culture. A total of 5 mL of the culture was inoculated into a 250 mL flask containing 50 mL of the pullulan producing media and then incubated at 28°C and 200 rpm for 7 days. After that, the fermentation broth was centrifuged at 14,000 × g for 30 min to remove the fungus cells. Then, 30 mL of the supernatant was taken to detect viscosity using a viscosimeter (DV3T, BROOKFIFLD, United States) with a 61-64# rotor at 30 rpm for 5 min. Collection and quantification of pullulan were performed according to the protocol described below. The pullulan producing media contained 140.0 g/L sucrose, 3.0 g/L yeast extract, 5.0 g/L K_2HPO_4, 0.2 g/L $MgSO_4{\cdot}7H_2O$, 0.01 g/L NaCl.

2.3 Phenotypic and molecular analysis of the fungal strain

The phenotypic analysis of the colonies formed on the PDA plates was performed in accordance with the methods described previously (Jiang et al., 2018; Liu et al., 2020). The genomic DNA of strain BL06 was extracted using a TIANamp Yeast DNA Kit (TIANGEN, Beijing, China). Amplification and sequencing of the internal transcribed spacer region (ITS) of the rRNA gene cluster were performed using the common primers ITS1 and ITS4 (Supplementary Table S2) according to the methods described by Ma et al. (2014). The obtained ITS sequence of strain BL06 was aligned using BLAST analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The phylogenetic tree was made in MEGA7.0 by the neighbor-joining method (Kumar et al., 2016; Liu et al., 2020).

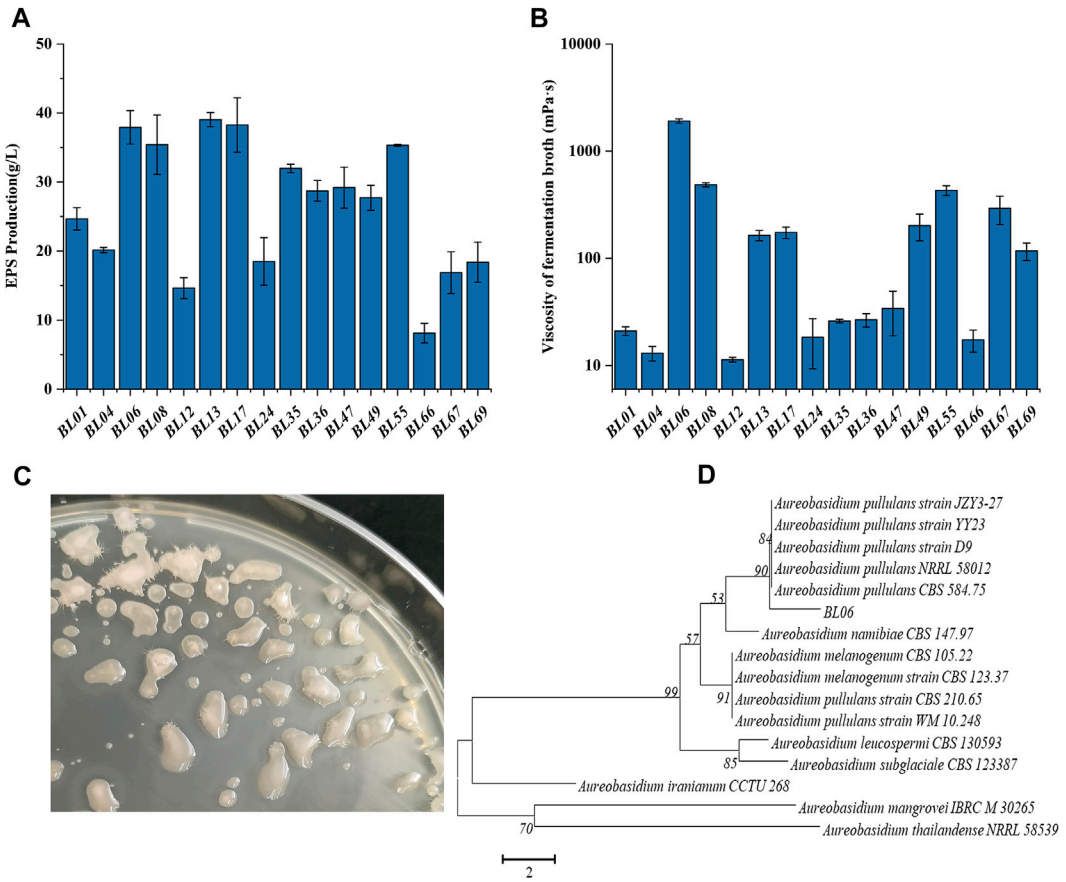


FIGURE 1
Screening and identification of wild strains for high production of EPS. **(A)** EPS production of the selected wild strains; **(B)** Viscosity of the fermentation broth of the selected wild strains; **(C)** Colonial morphology of strain BL06 on a PDA plate for 120 h of cultivation; **(D)** Phylogenetic tree of strain BL06 and other yeast-like fungal species relatives based on a neighbor-joining analysis of ITS sequences. All values are expressed as the means ± SDs (n = 3).

2.4 Purification and quantification of pullulan

The purification and quantification of pullulan were performed according to the methods reported by Ma et al. (2014) with slight modification. The fermentation broth was centrifuged at 4°C 14,000 × g for 10 min to remove the cell pellet. After that, twice the volume of ice-cold ethanol was added to the supernatant. The mixture was then incubated at 4°C for 24 h to precipitate pullulan. The precipitate was dissolved in deionized water at 80°C, followed by a repetitive ethanol precipitation procedure. The collected precipitates were lyophilized to quantify the dry weight of the purified pullulan from the fermentation products (Figure 2A).

2.5 Characterization of pullulan

To characterize the purified pullulan, high performance liquid chromatography (HPLC), fourier-transform infrared spectrometer (FT-IR) analysis (Jiang et al., 2019), right angle light scattering (RALS) (Liu et al., 2018), and high performance gel permeation chromatography-multiangle laser light scattering (HPGFC-MALLS) (Thomsen, 2020) were performed according to previous reports. The commercial pullulan (viscosity 100–180 mPa·s) and the experimental pullulan were dissolved in ddH_2O to obtain a 10 mg/mL pullulan solution, respectively. Taken 1 mL of the pullulan solution to be hydrolyzed by 2 U/mL pullulanase in 50 mM sodium acetate-acetic acid buffer (pH 4.5) at 60°C for 30 min. Then, the mixture was boiled for 10 min to terminate the hydrolysis reaction. The samples were then centrifuged at 14,000 × g for 10 min to collect the supernatant. The supernatant of hydrolysates was filtrated with 0.22 μm syringe filters before HPLC analysis. An Agilent 1260 HPLC with a refractive index detector (RID) and a Bio-Rad Aminex HPX-87P column was used in the analysis. The mobile phase was ultra-pure water with a flow rate of 0.6 mL/min at 65°C. In addition, maltotriose was used as an external standard for analysis of the hydrolysates. The lyophilized purified pullulan (4 mg) obtained in Section 2.4 was mixed with 120 mg of 95% potassium bromide powder and then desiccated overnight at 50°C under vacuum. FT-IR spectra were measured and recorded using potassium bromide pellets of the purified pullulan and the pullulan standard over an arrangement of 4,000–400 cm^{-1} at a rate of 16 scans with a resolution of 2 cm^{-2} using FT-IR (a Nicolet Nexus FTIR 470 spectrophotometer) (Jiang et al., 2019). The Mw of the purified pullulan was determined using RALS and HPGFC-MALLS, respectively. The lyophilized purified pullulan was dissolved in 0.1 N sodium nitrate solution. The sample was measured using RALS (Viscotek TDA305max, Malvern, UK) with an AGuard +1 × A6000 column. The mobile phase was 0.1 N sodium nitrate solution with a flow rate of 0.8 mL/min at 30°C (Liu et al., 2018). HPGFC-MALLS was specially used to determine the high Mw pullulan sample with a Mw above 3.0 × 10^6 Da. The lyophilized purified pullulan was dissolved in 0.1 N sodium nitrate solution to a final

concentration of 5 mg/mL. The sample was measured using an HPLC (Agilent 1260) with a wyatt multiangle light scattering detector and a Shodex OHpak SB-805 HQ column. The mobile phase was 0.1 N sodium nitrate solution with a flow rate of 0.5 mL/min at 40°C (Thomsen, 2020).

2.6 Gene knockout of *A. pullulans* BL06

The primers ΔAGSII-5F/5R, ΔAGSII-3F/3R, Δmel-5F/5R, Δmel-3F/3R, ΔPMAs-5F/5R, and ΔPMAs-3F/3R (Supplementary Table S2), used to amplify the homologous arm sequences of the pullulan synthase (AGSII) gene, polyketide synthase (PKS) gene, and polymalate synthase (PMAs) gene, were designed according to the genome sequences of *Aureobasidium melanogenum* P16, *Aureobasidium pullulans* As3.3984, and *Aureobasidium melanogenum* ATCC62921, respectively. The genome integration plasmids pSY018, pSY005, and pJQ046 (Supplementary Table S1 and Supplementary Figure S1) depending on homologous recombination were constructed by assembling the gene replacement boxes AGSII-5′arm-P*pgk*-NAT-T*polyA*-AGSII-3′arm, mel-5′arm-P*pgk*-NAT-T*polyA*-mel-3′arm, and PMAs-5′arm-P*tef*-HPT-T*tef*-PMAs-3′arm (Figure 3A,B) using the golden gate method. *Nat* and *Hpt*, the codon-optimized nourseothricin resistance gene and hygromycin B resistance gene, were expressed under the promoters of P*pgk* and P*tef* in the gene expression boxes, respectively (Figure 3A,B). The linear DNA was amplified by polymerase chain reaction (PCR) using ΔAGSII-5F/3R, Δmel-5F/5R, and ΔPMAs-5F/3R as primers from the templates of pSY018, pSY005, and pJQ046, respectively. The amplified DNAs (1–2 μg) were transformed into protoplasts of *A. pullulans* BL06 through the electro-transformation method. The recombinant strains were screened on YPD plates containing NAT or HPT at a concentration of 50.0 μg/mL. The positive engineered strains were determined by PCR and sequencing of the gene integration site on the genome. The double-gene knockout strain *A. pullulans* BL06ΔPMAsΔmel was constructed by two transformations in succession and screened for both NAT and HPT resistance.

2.7 Identification of polymalic acid content

The engineered strain *A. pullulans* BL06 was cultivated in a pullulan producing medium at 28°C and 200 rpm for 120 h. After centrifugation at 14,000 × g for 10 min, 1 mL of fermentation supernatant was mixed with 1 mL of sulfuric acid solution (1 M) and placed at 90°C for 12 h. The reaction supernatant was centrifuged at 14,000 × g for 10 min to be filtrated with 0.22 μm syringe filters before HPLC analysis. The Bio-Rad Aminex HPX-87H column was used to determine the concentration of polymalic acid under an ultraviolet detector at 210 nm. The mobile phase was 5 mM sulfuric acid solution with a flow rate of 0.6 mL/min at 65°C.

2.8 Fermentation process of *A. pullulans* strains for pullulan production in a 5 L bioreactor

The pullulan production of *A. pullulans* BL06 and *A. pullulans* BL06ΔPMAs were conducted *via* batch fermentation in a 5 L bioreactor for 120 h. The preserved strains were streaked on YPD plates and cultured at 28°C for 48–72 h. The single colonies were picked up and inoculated in 5 mL YPD medium. After cultivation at 28°C and 200 rpm for 20 h, the culture broth was transferred into 300 mL fresh YPD medium in a shake flask for another 16 h cultivation at the same condition. The obtained 300 mL of seed culture was then added into a 5 L bioreactor at 10% (V/V) of the dose. The total liquid content was 60% (V/V) of the bioreactor The cultivation was performed at 28 ± 2°C with a stirring speed of 500 rpm. Once the dissolved oxygen was reduced to 20%, the stirring speed was increased to 800 rpm and the ventilation was maintained at 3 vvm. After 36 h fermentation, 800 mL of sucrose liquid (140 g/L) was feeding at 100 mL/h. During the fermentation, 50 mL of the culture was harvested at 12 h interval for product detecting. The fermentation medium components in the bioreactor included 11.6% (w/v) sucrose, 0.28% (w/v) yeast extract powder, 0.07% (w/v) ammonium sulfate, 0.45% (w/v) K_2HPO_4, 0.02% (w/v) $MgSO_4{\cdot}7H_2O$, 0.09% (w/v) NaCl.

2.9 Application of the novel moderate Mw pullulan in food preservation

The solution (30 g/L) of purified pullulan produced by *A. pullulans* BL06 ΔPMAs was spread evenly to the surface of celery cabbages and mangos, respectively. The original weights of the selected samples were similar. After laying at room temperature (25°C–35°C) with a relative humidity of 68%–75% for different times (1 day, 2 days, 3 days, to 15 days), the surface state of the samples was observed and the weight loss of the samples was determined, respectively. The samples with no pullulan spread were used as controls.

2.10 Analytic methods and data availability

The statistical analysis was carried out using one-way analysis of variance followed by Duncan's multiple comparison tests. p values < 0.05 were considered statistically significant. The results are presented as the mean ± the standard deviation (SD) for a replication of $n = 3$. The ITS sequence of *A. pullulans* BL06 was deposited in the GenBank database with the accession number OP810667. The gene sequences of AGSII, mel, and PMAs were based on the records in the GenBank database with accession numbers MH917125.1, KT429644, and MN551082, respectively.

3 Results and discussion

3.1 A high-yield strain of high-Mw pullulan was screened from the environment

Through morphologically screening of the colonies on the PDA plate, 16 strains forming with smooth, moist, and yeast-like properties, as shown in Figure 1C, were chosen for further evaluation of their ability to produce extracellular polysaccharides (EPS). Each of the strains was incubated in a pullulan producing media at 28°C and 200 rpm for 7 days. As shown in Figure 1A, all 16 strains showed the ability to produce EPS. Among those, the strains BL06, BL13 and BL17 produced the relatively highest EPS amount (≥37 g/L). At the same time, the rheological behavior of the extracellular polysaccharides produced by the 16 strains was determined. As the results showed in Figure 1B, the viscosity of the EPS produced by most strains was lower than 1,000 mPa·s, and only strain BL06 showed the significantly highest viscosity of 1,912 mPa·s at 37.9 g/L EPS. However, similar amounts (39 g/L and 38.3 g/L) of EPS production of

10.3389/fbioe.2023.1131875

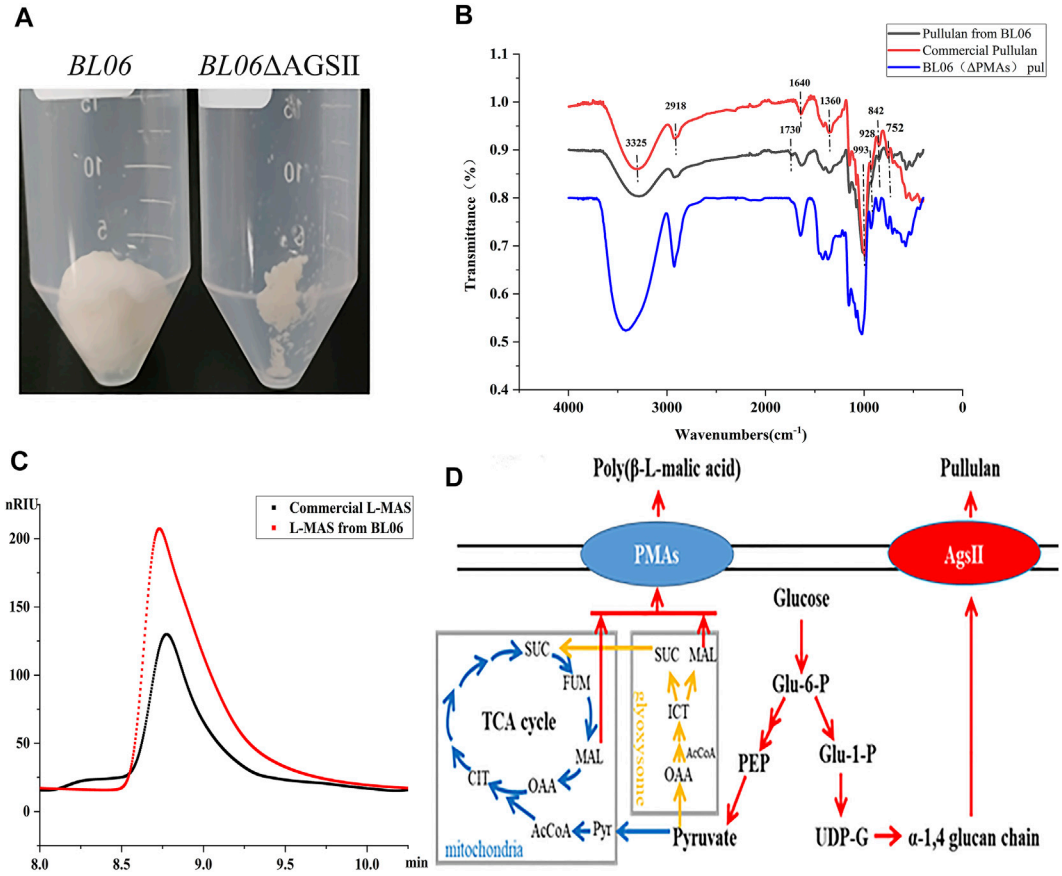


FIGURE 2
Identification and properties of the EPS of *A. pullulans* BL06. **(A)** The purified EPS produced by *A. pullulans* BL06 and *A. pullulans* BL06ΔAGSII; **(B)** FTIR spectra of the purified EPS and standard pullulan; **(C)** HPLC analysis of the hydrolysate of the purified EPS produced by *A. pullulans* BL06ΔAGSII and L-malic acid standard. **(D)** Synthetic pathways of pullulan and polymalic acid in *A. pullulans*.

strains BL13 and BL17 showed viscosities of only 163 mPa·s and 174 mPa·s, respectively (Figures 1A, B). Thus, the remarkably higher viscosity makes us interested in further identifying the properties of the EPS produced by strain BL06. Before identifying the EPS properties of strain BL06, the ITS sequence of strain BL06 was determined and uploaded to the GenBank database. After the search for similarities between ITSs of the isolate strain BL06 and those of the type strains in the National Center for Biotechnology Information (NCBI) database, a phylogenetic tree of ITSs that showed higher than 98% identity with that of strain BL06 was constructed, as shown in Figure 1D. The results showed that many phylogenetically related *Aureobasidium spp.* were similar to strain BL06 and the most closely related strain was *A. pullulans* (Figure 1D). So the newly isolated strain BL06 was identified as *A. pullulans* BL06 in this study. According to previous reports, *Aureobasidium pullulans* species are known to be hosts for high level production of pullulan (Ma et al., 2014; Liu et al., 2021).

3.2 Production properties of the yeast-like fungus Aureobasidium pullulans BL06

In order to further identify the high viscosity EPS produced by *A. pullulans* BL06. The EPS in 3 L of fermentation broth was purified by the alcohol precipitation method to obtain 90 g white dry powders (Figure 2A). The purified EPS was analyzed by FT-IR. As shown in Figure 2B, in comparison with the commercial pullulan standard as a control, except for a peak at 1730 cm^{-1}, all of the characteristic peaks of EPS from strain BL06 were consistent with pullulan. The intense absorbance peak appears at 3,325 cm^{-1}, which was the unique characteristic of the presence of repeating units of OH groups in sugars. The peak at 2,918 cm^{-1} showed the presence of C-H stretching. The absorption at 842 cm^{-1} was attributed to the presence of α-D-glucopyranoside bonds whereas a band at 752 cm^{-1} revealed the presence of α-(1, 4)-D-glucosidic bonds, and a strong absorption at 993 cm^{-1} was characteristic of α-(1, 6)-D-glucosidic bonds. The strong peaks at approximately 1640 cm^{-1} and 1360 cm^{-1} were characteristic of glycosidic linkage O-C-O bonds and C-O-H bending, respectively. These absorption patterns at various stretching frequencies confirmed that the purified EPS produced by *A. pullulans* BL06 had all the pullulan-like peaks and was identified as pullulan. The Mw of the purified pullulan was 3.3×10^6 Da determined by the HPGPC-MALLS method (Supplementary Figure S2). It is significantly higher than that of the previously reported level (Jiang et al., 2019; Singh et al., 2019). This result could also explain why the exopolysaccharides produced by strain 06 showed remarkably higher viscosity at similar EPS concentration compared with other strains (Figures 1A, B). Because pullulan as a straight chain polymer, the property of viscosity was mainly affected by its Mw and concentration (Singh et al., 2021). Thus,

10.3389/fbioe.2023.1131875

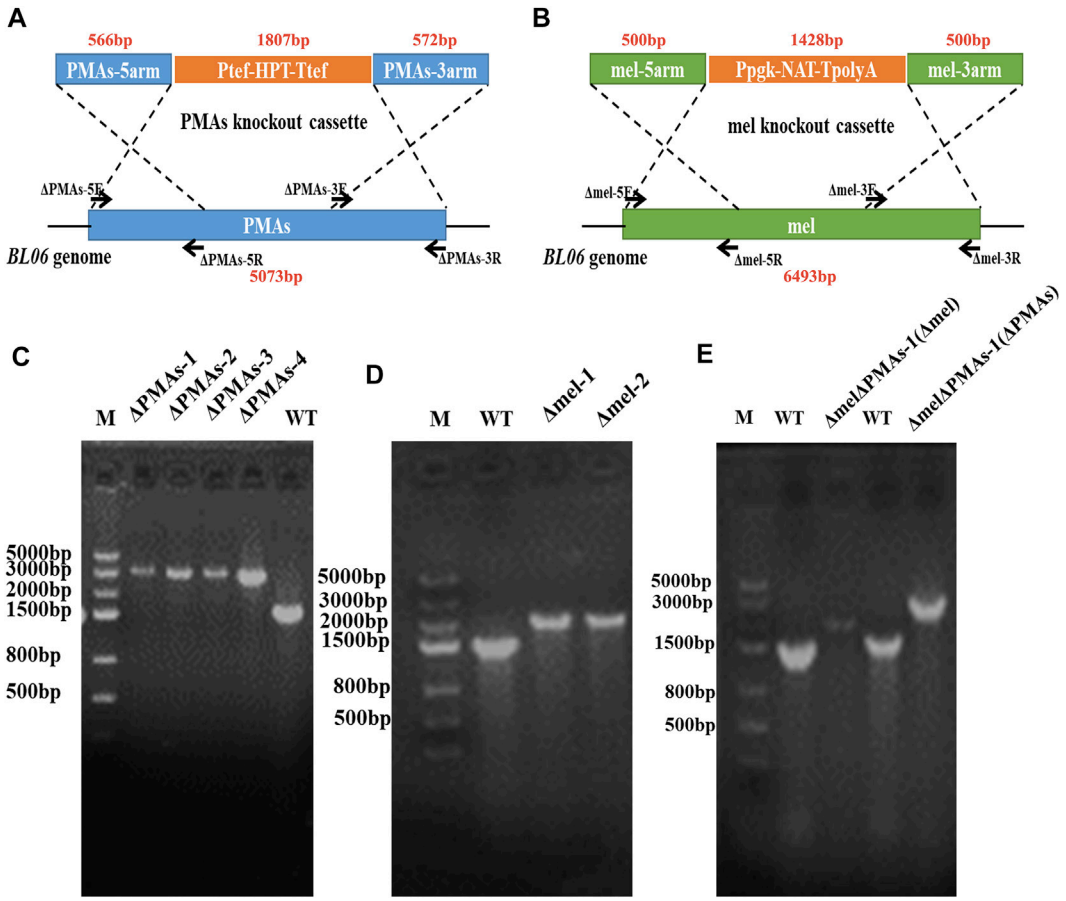


FIGURE 3
Construction of engineered strains for gene knockout. **(A)** Schematic diagram of gene knockout of PMAs in the genome of *A. pullulans* BL06; **(B)** Schematic diagram of gene knockout of melanin in the genome of *A. pullulans* BL06; **(C)** Identification of the positive transformants for *A. pullulans* BL06ΔPMAs by PCR; **(D)** Identification of the positive transformants for *A. pullulans* BL06Δmel by PCR; **(E)** Identification of the positive transformants for *A. pullulans* BL06ΔPMAsΔmel by PCR.

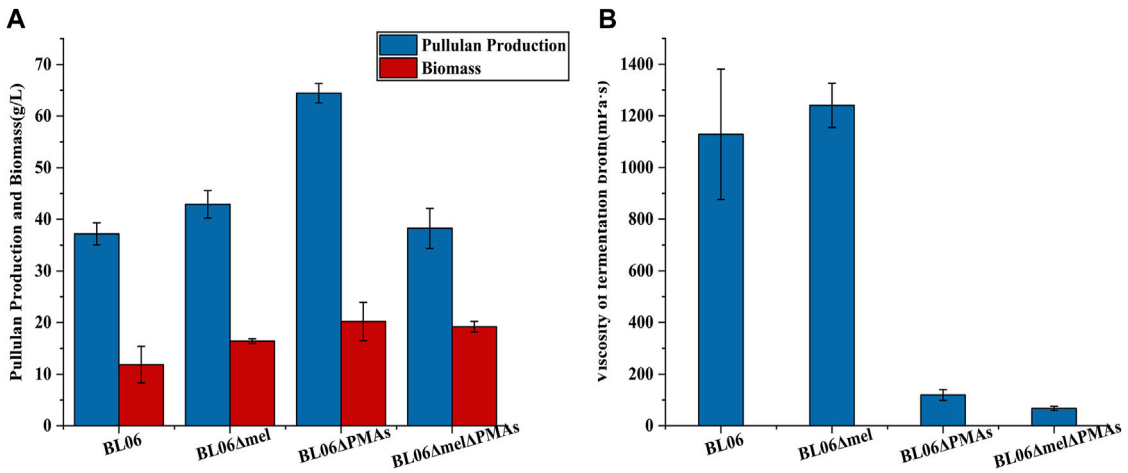


FIGURE 4
Pullulan production and viscosity of fermentation broth of the engineered strains. **(A)** Pullulan production and biomass of the engineered strains after 120 h fermentation; **(B)** Viscosity of fermentation broth of the engineered strains after 120 h fermentation. All values are expressed as the means ± SDs (n = 3).

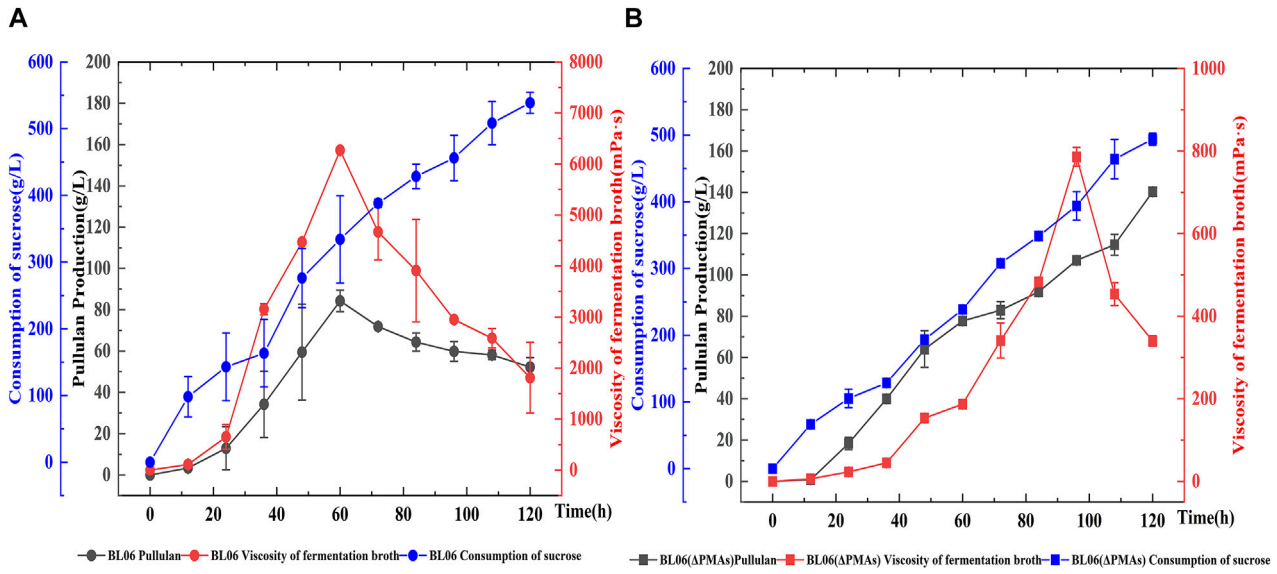


FIGURE 5
Batch feeding fermentation of *A. pullulans* BL06 **(A)** and *A. pullulans* BL06ΔPMAs **(B)** in a 5 L bioreactor. The black solid circle and solid square indicate the pullulan production by *A. pullulans* BL06 and *A. pullulans* BL06ΔPMAs respectively; the red solid circle and solid square indicate the viscosity of the fermentation broth of *A. pullulans* BL06 and *A. pullulans* BL06ΔPMAs respectively; the blue solid circle and solid square indicate the consumption of sucrose by *A. pullulans* BL06 and *A. pullulans* BL06ΔPMAs respectively. All values are expressed as the means ± SDs (n = 3).

the ability to produce high-Mw pullulan endowed *A. pullulans* BL06 with important value in scientific research and industrial applications.

To determine the purity of extracellular pullulan produced by *A. pullulans* BL06, the pullulan synthase (AGSII) gene (Figure 2D) was knocked out from its genome. Through weighting the purified extracellular alcohol precipitation products (Figure 2A) of *A. pullulans* BL06 ΔAGSII, the pullulan purity produced by *A. pullulans* BL06 was calculated to be no less than 96.9%. Besides, based on the specific peak at 1730 cm^{-1} of samples produced by *A. pullulans* BL06 revealing the presence of -C = O (Figure 2B) and *A. pullulans* strains were also reported to yield a high level of polymalic acids (Figure 2D) (Zeng et al., 2020), the impurities were estimated as polymalic acids. To further prove the estimation, the alcohol precipitation products of *A. pullulans* BL06 (Figure 2A) were identified by HPLC after acid hydrolysis. The results showed that the peak of the acid hydrolysis products was consistent with the characteristic peak of the L-malic acid standard (Figure 2C). Though quantitative calculation according to the methods described in Section 2.7, *A. pullulans* BL06 screened in this study produced a proportion (3%) of polymalic acids during the production of the high-Mw pullulan.

3.3 Strain modification of A. pullulans BL06 for the production of pullulan

The byproducts (polymalic acid) produced by *A. pullulans* BL06 will affect the purification of pullulan (Figure 2A) from the fermentation broth to obstruct its application in various fields. Besides, the synthesis of polymalic acid also competed for the substrate glucose-6-phosphate (Figure 2D) to reduce the amount of pullulan in *A. pullulans* BL06. Thus, an engineered strain *A. pullulans* BL06ΔPMAs was constructed through knocking out the encoding gene for the polymalic acid synthase (PMAs) in the genome of *A. pullulans* BL06. Moreover, according to previous reports, *A. pullulans* species strains usually synthesize melanin, thereby being named "black yeast" (Liu et al., 2020). While the melanin production is well known as an obstacle to pullulan industrial production because it increases the cost of pullulan purification (Liu et al., 2020). Therefore, a melanin deficiency strain *A. pullulans* BL06Δmel and a double gene knockout strain *A. pullulans* BL06ΔmelΔPMAs were also constructed in this study. All three engineered strains were verified to have the target genes knocked out correctly (Figures 3C–E). Compare with the titer of pullulan produced by wild-type *A. pullulans* BL06, all three engineered strains showed increase in the titer of pullulan (Figure 4A). In particular, *A. pullulans* BL06ΔPMAs produced pullulan at the highest yield level of 64.43 g/L after 120 h of fermentation, which was 1.7 times higher than that of *A. pullulans* BL06 (Figure 4A). *A. pullulans* BL06Δmel and *A. pullulans* BL06ΔmelΔPMAs produced slightly higher amounts of pullulan than *A. pullulans* BL06, while taking the growth enhancement of the two strains into account indicated that melanin deficiency could fail to increase pullulan production and even reduce it (Figure 4A). As a previous report, the pullulan biosynthetic genes *upt*, *pgm*, *ugp*, and *pul* were downregulated, while the negative regulatory gene (*creA*) of pullulan synthesis was upregulated by melanin deficiency (Liu et al., 2021). In this study, the melanin deficiency regulation effect was more significant in *A. pullulans* BL06ΔPMAs (Figure 4A). Besides, melanin deficiency did not affect the viscosity of extracellular pullulan products, while the viscosity of pullulan (119 mPa·s) was decreased remarkably due to knocking out PMAs. Based on determining the Mw (1.3×10^5) of the extracellular pullulan produced by *A. pullulans* BL06ΔPMAs (Supplementary Figure S3), the decrease in viscosity was considered as be caused by the reduction of Mw of pullulan. Thus, a novel engineered strain, *A. pullulans* BL06ΔPMAs, showed a high level

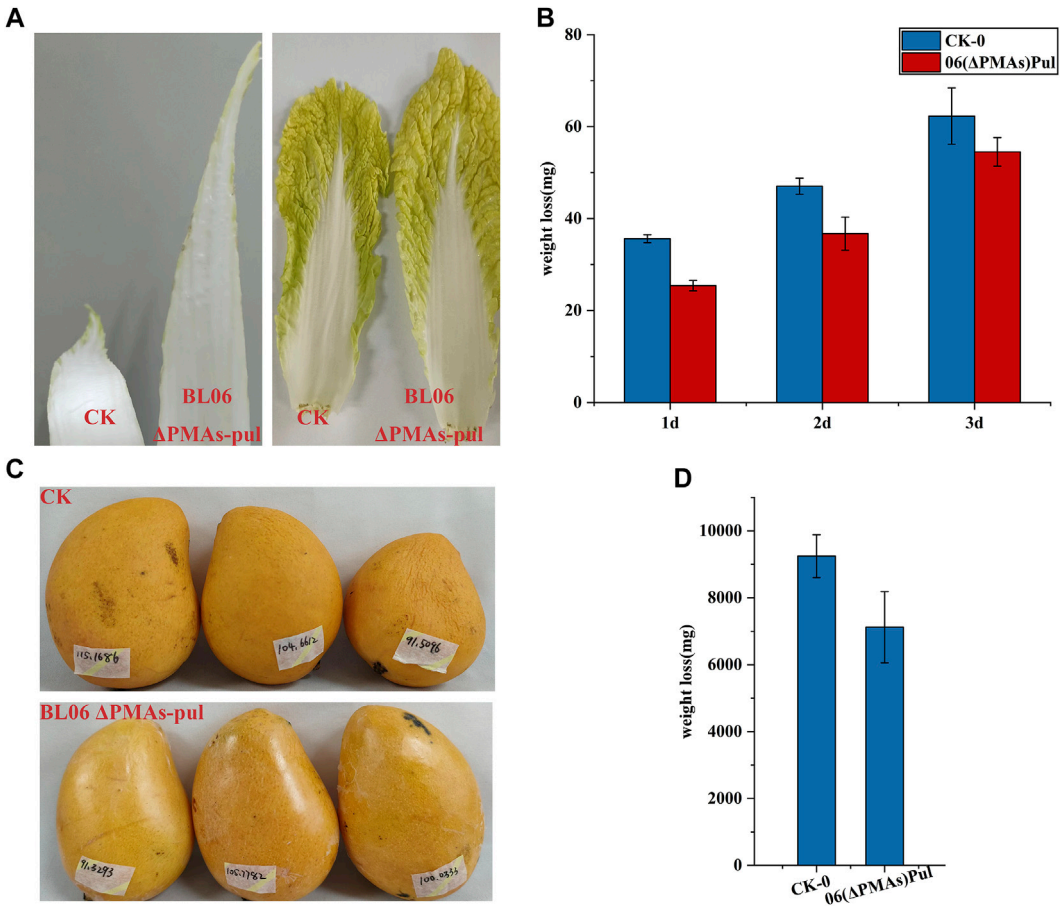


FIGURE 6
Food coating application of the moderate Mw pullulan films. **(A)** Comparison of apparent states of celery cabbages smeared with pullulan films or not; **(B)** Comparison of apparent states of mangos smeared with pullulan films or not; **(C)** Comparison of weight loss of celery cabbages smeared with pullulan films or not; **(D)** Comparison of weight loss of mangos smeared with pullulan films or not; All values are expressed as the means ± SDs (n = 3).

of extracellular moderate Mw pullulan was realized in this study. Through the results, we also conclude that regulating the expression of PMAs in *A. pullulans* BL06 may effectively change the Mw and titer of pullulan in *A. pullulans* BL06. The specific regulatory mechanism needs to be further revealed.

3.4 The scale-up fermentation of the pullulan production strains in a 5 L bioreactor

A batch feeding fermentation of *A. pullulans* BL06 and *A. pullulans* BL06ΔPMAs in a 5 L bioreactor was performed to produce high Mw pullulan and moderate Mw pullulan, respectively. The highest yield (84.3 g/L) of extracellular high-Mw pullulan with a high viscosity of 6,270 mPa·s was obtained for *A. pullulans* BL06 at 60 h of fermentation (Figure 5). After 60 h, the high-Mw pullulan was degraded gradually associating the viscosity decreases (Figure 5). However, the yield of the moderate-Mw pullulan produced by *A. pullulans* BL06ΔPMAs gradually increased during fermentation. At 120 h of fermentation, the highest yield of the moderate-Mw pullulan was up to 140.2 g/L with a viscosity of 340 mPa·s (Figure 5). It is the highest level of pullulan to date. Besides, the sucrose consumption profiles of the two strains were similar, while *A. pullulans* BL06ΔPMAs consumed relatively less sucrose substrate (Figure 5). Thus, the novel engineered strain *A. pullulans* BL06ΔPMAs could vastly reduce the production cost and expand the application scope and potential of pullulan.

3.5 The application effect of novel pullulan in food preservation

Pullulan is an edible polymer and has been certified to be harmless for usage in food products by food safety regulations in many countries (Singh et al., 2019). Pullulan solutions can form clear films or coatings that are oxygen impermeable, oil resistant and have good mechanical properties (Singh et al., 2019). Owing to their properties, pullulan films have the potential to be used as versatile and novel packaging materials. Combined with the high production level of the moderate-Mw pullulan obtained in this study will significantly reduce its application cost. Thus, the moderate-Mw pullulan produced by *A. pullulans* BL06ΔPMAs was used as a protective film smearing evenly to the surface of celery cabbages and mangos, respectively. After 3 days and 15 days of laying at room temperature environment (25–35°C and 68%–75% humidity), the celery cabbages and mangos coated with pullulan remained fresh

and had a good appearance, while the control group was shrunken and wrinkled (Figures 6A, B). Further weighting of the samples showed that the pullulan coating reduced the weight loss of celery cabbages and mangos by 12.5% and 22.0%, respectively, compared with the controls (Figure 6C, D). The results indicated that the moderate-Mw pullulan developed in this study could be used as a food coating to enhance the shelf-life of vegetables and fruits, which protects them from dehydration and spoilage.

4 Conclusion

In this study, a novel strain *A. pullulans* BL06 with a high production level for high Mw pullulan was firstly screened and identified. By regulating the key synthesis pathway of pullulan impurities, we found that knocking out the key genes not only affected the yield and purity of extracellular pullulan but also affected the Mw of pullulan products. The engineered strain *A. pullulans* BL06ΔPMAs with the highest enhancement in pullulan production was discovered to produce the extracellular pullulan with a mean moderate Mw. After scale-up fermentation in a 5 L bioreactor, the highest production of the high-Mw (3.3×10^6 Da) pullulan was 84.3 g/L with a viscosity of 6,270 mPa·s produced by *A. pullulans* BL06, and the highest production of the moderate-Mw (1.3×10^5 Da) pullulan was up to 140.2 g/L with a viscosity of 340 mPa·s produced by *A. pullulans* BL06ΔPMAs, respectively. Through the two strains and the key regulated genes obtained in this study, pullulan products with various Mw ranges will be synthesized efficiently in further work. Besides, the high production level of the moderate-Mw pullulan obtained in this study was certified for use as a food coating to enhance the shelf-life of vegetables and fruits. However, the application of the various pullulans developed in this study is definitely not limited to food preservation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

SC contributed to methodology, validation, formal analysis, and investigation. HZ wrote the original draft and performed writing-review and editing. JG contributed to methodology and investigation. WB organized conceptualization, project administration, supervision, date curation, and funding acquisition. HS was responsible for supervision. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFC2103200), and the Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-CXRC-010).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1131875/full#supplementary-material

References

An, J. M., Shahriar, S. M. S., Hasan, M. N., Cho, S., and Lee, Y. K. (2021). Carboxymethyl cellulose, pluronic, and pullulan-based compositions efficiently enhance antiadhesion and tissue regeneration properties without using any drug molecules. *ACS Appl. Mater Interfaces* 13 (14), 15992–16006. doi:10.1021/acsami.0c21938

Chen, L., Chi, Z., Liu, G. L., Xue, S. J., Wang, Z. P., Hu, Z., et al. (2019). Improved pullulan production by a mutant of Aureobasidium melanogenum TN3-1 from a natural honey and capsule shell preparation. *Int. J. Biol. Macromol.* 141, 268–277. doi:10.1016/j.ijbiomac.2019.08.264

Chen, L., Wei, X., Liu, G. L., Hu, Z., Chi, Z. M., and Chi, Z. (2020). Glycerol, trehalose and vacuoles had relations to pullulan synthesis and osmotic tolerance by the whole genome duplicated strain Aureobasidium melanogenum TN3-1 isolated from natural honey. *Int. J. Biol. Macromol.* 165, 131–140. doi:10.1016/j.ijbiomac.2020.09.149

Chen, X., Wang, Y., He, C. Y., Wang, G. L., Zhang, G. C., Wang, C. L., et al. (2021). Improved production of beta-glucan by a T-DNA-based mutant of Aureobasidium pullulans. *Appl. Microbiol. Biotechnol.* 105 (18), 6887–6898. doi:10.1007/s00253-021-11538-x

Ding, Y., Jiang, F., Chen, L., Lyu, W., Chi, Z., Liu, C., et al. (2020). An alternative hard capsule prepared with the high molecular weight pullulan and gellan: Processing, characterization, and *in vitro* drug release. *Carbohydr. Polym.* 237, 116172. doi:10.1016/j.carbpol.2020.116172

Duan, Y. M., Li, K. Y., Wang, H. W., Wu, T., Zhao, Y. F., Li, H. Y., et al. (2020). Preparation and evaluation of curcumin grafted hyaluronic acid modified pullulan polymers as a functional wound dressing material. *Carbohydr. Polym.* 238, 116195. doi:10.1016/j.carbpol.2020.116195

Feng, Z., Chen, S., Ahmad, A., Chen, L., and Bai, W. (2022). Ultra-high molecular weight pullulan-based material with high deformability and shape-memory properties. *Carbohydr. Polym.* 295, 119836. doi:10.1016/j.carbpol.2022.119836

Ghorbani, F., Zamanian, A., Behnamghader, A., and Joupari, M. D. (2020). Bioactive and biostable hyaluronic acid-pullulan dermal hydrogels incorporated with biomimetic hydroxyapatite spheres. *Mater. Sci. Eng. C-Materials Biol. Appl.* 112, 110906. doi:10.1016/j.msec.2020.110906

Jiang, H., Chen, T. J., Chi, Z., Hu, Z., Liu, G. L., Sun, Y., et al. (2019). Macromolecular pullulan produced by Aureobasidium melanogenum 13-2 isolated from the Taklimakan desert and its crucial roles in resistance to the stress treatments. *Int. J. Biol. Macromol.* 135, 429–436. doi:10.1016/j.ijbiomac.2019.05.190

Jiang, H., Xue, S. J., Li, Y. F., Liu, G. L., Chi, Z. M., Hu, Z., et al. (2018). Efficient transformation of sucrose into high pullulan concentrations by Aureobasidium melanogenum TN1-2 isolated from a natural honey. *Food Chem.* 257, 29–35. doi:10.1016/j.foodchem.2018.03.003

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi:10.1093/molbev/msw054

Liu, F., Zhang, J., Zhang, L., Diao, M., Ling, P., and Wang, F. (2021). Correlation between the synthesis of pullulan and melanin in Aureobasidium pullulans. *Int. J. Biol. Macromol.* 177, 252–260. doi:10.1016/j.ijbiomac.2021.02.108

Liu, G., Zhao, X., Chen, C., Chi, Z., Zhang, Y., Cui, Q., et al. (2020). Robust production of pigment-free pullulan from lignocellulosic hydrolysate by a new fungus co-utilizing glucose and xylose. *Carbohydr. Polym.* 241, 116400. doi:10.1016/j.carbpol.2020.116400

Liu, N. N., Chi, Z., Liu, G. L., Chen, T. J., Jiang, H., Hu, Z., et al. (2018). α-Amylase, glucoamylase and isopullulanase determine molecular weight of pullulan produced by Aureobasidium melanogenum P16. *Int. J. Biol. Macromol.* 117, 727–734. doi:10.1016/j.ijbiomac.2018.05.235

Ma, Z. C., Fu, W. J., Liu, G. L., Wang, Z. P., and Chi, Z. M. (2014). High-level pullulan production by Aureobasidium pullulans var. melanogenium P16 isolated from mangrove system. *Appl. Microbiol. Biotechnol.* 98 (11), 4865–4873. doi:10.1007/s00253-014-5554-5

Moraes, F. C., Antunes, J. C., Forero Ramirez, L. M., Aprile, P., Franck, G., Chauvierre, C., et al. (2020). Synthesis of cationic quaternized pullulan derivatives for miRNA delivery. *Int. J. Pharm.* 577, 119041. doi:10.1016/j.ijpharm.2020.119041

Nishimura, T., Shishi, S., Sasaki, Y., and Akiyoshi, K. (2020). Thermoresponsive polysaccharide graft polymer vesicles with tunable size and structural memory. *J. Am. Chem. Soc.* 142 (27), 11784–11790. doi:10.1021/jacs.0c02290

Shah, S. A., Sohail, M., Minhas, M. U., Khan, S., Hussain, Z., Mahmood, A., et al. (2021). Curcumin-laden hyaluronic acid-co-Pullulan-based biomaterials as a potential platform to synergistically enhance the diabetic wound repair. *Int. J. Biol. Macromol.* 185, 350–368. doi:10.1016/j.ijbiomac.2021.06.119

Singh, R. S., Kaur, N., Hassan, M., and Kennedy, J. F. (2021). Pullulan in biomedical research and development - a review. *Int. J. Biol. Macromol.* 166, 694–706. doi:10.1016/j.ijbiomac.2020.10.227

Singh, R. S., Kaur, N., and Kennedy, J. F. (2019). Pullulan production from agro-industrial waste and its applications in food industry: A review. *Carbohydr. Polym.* 217, 46–57. doi:10.1016/j.carbpol.2019.04.050

Sugumaran, K. R., and Ponnusami, V. (2017). Review on production, downstream processing and characterization of microbial pullulan. *Carbohydr. Polym.* 173, 573–591. doi:10.1016/j.carbpol.2017.06.022

Thomsen, M. (2020). Determination of the molecular mass of membrane proteins using size-exclusion chromatography with multiangle laser light scattering (SEC-MALLS). *Methods Mol. Biol.* 2168, 263–269. doi:10.1007/978-1-0716-0724-4_12

Wei, X., Liu, G. L., Jia, S. L., Chi, Z., Hu, Z., and Chi, Z. M. (2021). Pullulan biosynthesis and its regulation in Aureobasidium spp. *Carbohydr. Polym.* 251, 117076. doi:10.1016/j.carbpol.2020.117076

Xue, S. J., Chen, L., Jiang, H., Liu, G. L., Chi, Z. M., Hu, Z., et al. (2019). High pullulan biosynthesis from high concentration of glucose by a hyperosmotic resistant, yeast-like fungal strain isolated from a natural comb-honey. *Food Chem.* 286, 123–128. doi:10.1016/j.foodchem.2019.01.206

Zeng, W., Zhang, B., Jiang, L., Liu, Y., Ding, S., Chen, G., et al. (2020). Poly(malic acid) production from liquefied corn starch by simultaneous saccharification and fermentation with a novel isolated Aureobasidium pullulans GXL-1 strain and its techno-economic analysis. *Bioresour. Technol.* 304, 122990. doi:10.1016/j.biortech.2020.122990

Zhang, Z., Lu, Y., Chi, Z., Liu, G. L., Jiang, H., Hu, Z., et al. (2019). Genome editing of different strains of Aureobasidium melanogenum using an efficient Cre/loxp site-specific recombination system. *Fungal Biol.* 123 (10), 723–731. doi:10.1016/j.funbio.2019.06.001

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 06 February 2023
DOI 10.3389/fbioe.2023.1118702

OPEN ACCESS

EDITED BY
Yinjie Tang,
Washington University in St. Louis,
United States

REVIEWED BY
Jinjin Diao,
Washington University in St. Louis,
United States
Shihui Yang,
Hubei University, China

*CORRESPONDENCE
Pablo Carbonell,
pablo.carbonell@upv.es

SPECIALTY SECTION
This article was submitted
to Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

RECEIVED 07 December 2022
ACCEPTED 26 January 2023
PUBLISHED 06 February 2023

CITATION
Tellechea-Luzardo J, Stiebritz MT and
Carbonell P (2023), Transcription factor-
based biosensors for screening and
dynamic regulation.
Front. Bioeng. Biotechnol. 11:1118702.
doi: 10.3389/fbioe.2023.1118702

Transcription factor-based biosensors for screening and dynamic regulation

Jonathan Tellechea-Luzardo[1], Martin T. Stiebritz[1] and Pablo Carbonell[1,2]*

[1]Institute of Industrial Control Systems and Computing (AI2), Universitat Politècnica de València (UPV), Valencia, Spain, [2]Institute for Integrative Systems Biology I2SysBio, Universitat de València-CSIC, Paterna, Spain

Advances in synthetic biology and genetic engineering are bringing into the spotlight a wide range of bio-based applications that demand better sensing and control of biological behaviours. Transcription factor (TF)-based biosensors are promising tools that can be used to detect several types of chemical compounds and elicit a response according to the desired application. However, the wider use of this type of device is still hindered by several challenges, which can be addressed by increasing the current metabolite-activated transcription factor knowledge base, developing better methods to identify new transcription factors, and improving the overall workflow for the design of novel biosensor circuits. These improvements are particularly important in the bioproduction field, where researchers need better biosensor-based approaches for screening production-strains and precise dynamic regulation strategies. In this work, we summarize what is currently known about transcription factor-based biosensors, discuss recent experimental and computational approaches targeted at their modification and improvement, and suggest possible future research directions based on two applications: bioproduction screening and dynamic regulation of genetic circuits.

KEYWORDS

allosteric transcription factors, biosensors, screening, dynamic regulation, metabolic engineering

1 Introduction

Biosensors are biological devices combining two essential components: a sensing component that detects a particular input—typically the presence of a chemical—and a reporter that produces a measurable output after receiving the signal transduced by the sensing component. Whole-cell biosensors use biochemical transformations inside living cells to detect and react to different inputs (Fernandez-López et al., 2015).

One important class of whole-cell biosensors are those based on transcription factors (TF). TFs are proteins that can control the expression of genes by binding to specific DNA sequences. Some TFs are triggered after binding to a metabolite or external compound (known as allosteric transcription factors, aTFs). Once activated, a conformational change in the TF makes itself release from or attach to the DNA sequence upstream of the target gene, thereby activating or repressing its expression. TFs can be assembled together with other DNA parts commonly used in synthetic biology, such as promoters, ribosome binding sites (RBSs), terminators and reporter genes, to create TF-based biosensor circuits. These genetic devices can thus be used to sense and react to a range of intracellular or environmental ligand concentrations (De Paepe et al., 2018). Even

TABLE 1 Name, description, bibliographic reference, and web link to some of the most important TF databases. The description information was directly taken from the website.

Databases	
Name and description	References
P2TF (Predicted Prokaryotic Transcription Factors): an integrated and comprehensive database of TF proteins, which contains a compilation of the TF genes within completely sequenced genomes and metagenomes	Ortet et al. (2012)
JASPAR: an open-access database of curated, non-redundant transcription factor (TF) binding profiles stored as position frequency matrices (PFMs) and TF flexible models (TFFMs) for TFs across multiple species in six taxonomic groups	Castro-Mondragon et al. (2022)
TF2DNA: database provides comprehensive information about transcription factor binding motifs and their regulated genes for five model organisms and humans	Pujato et al. (2014)
GRASSIUS: Divided in GrassTFDB which provides a comprehensive collection of transcription factors from maize, sugarcane, sorghum and rice and GrassCoRegDB which provides a collection of proteins that are transcriptional regulatory factors but do not bind DNA in a sequence specific fashion	Yilmaz et al. (2009)
RegulonDB: the primary database on transcriptional regulation in *Escherichia coli* K-12	Santos-Zavaleta et al. (2019)
SM-TF database: collects available 3D structures of small molecule-transcription factor complexes from Protein Data Bank (PDB)	Xu et al. (2016)
CollecTF: a database of transcription factor binding sites (TFBS) in the Bacteria domain	Kılıç et al. (2014)
AnimalTFDB3: a comprehensive database including classification and annotation of genome-wide transcription factors (TFs), and transcription cofactors in 97 animal genomes	Hu et al. (2019)
PlantTFDB: Plant Transcription Factor Database	Jin et al. (2017)
RegPrecise: a database for capturing, visualisation and analysis of transcription factor regulons that were reconstructed by the comparative genomic approach in a wide variety of prokaryotic genomes	Novichkov et al. (2013)
SigMol: a repertoire of Quorum Sensing Signalling Molecules in Prokaryotes	Rajput et al. (2016)
Bionemo: stores manually curated information about proteins and genes directly implicated in the Biodegradation metabolism	Carbajosa et al. (2009)
PRODORIC: a comprehensive database about gene regulation and gene expression in prokaryotes. It includes a manually curated and unique collection of transcription factor binding sites	Dudek and Jahn (2022)
Tools	
footprintDB: predicts transcription factors which bind a specific DNA site or motif and DNA motifs or sites likely to be recognized by a specific DNA-binding protein	Sebastian and Contreras-Moreira (2014)
CiiDER: a user-friendly tool for predicting and analysing transcription factor binding sites, designed with biologists in mind	Gearing et al. (2019)
BART (Binding Analysis for Regulation of Transcription): a bioinformatics tool for predicting functional transcriptional regulators (TRs)	Wang et al. (2018)
PROMO: a virtual laboratory for the identification of putative transcription factor binding sites (TFBS) in DNA sequences from a species or groups of species of interest	Messeguer et al. (2002)
DeepTFactor: a deep learning-based tool for the prediction of transcription factors	Kim et al. (2021)

though allosteric transcription factors are suitable building blocks for the design of biosensors, they might require prior optimization or changes to their ligand specificity. The literature contains several examples in which sophisticated directed-evolution strategies were applied to this end (Wu et al., 2017; Machado et al., 2019; Berepiki et al., 2020; Snoek et al., 2020), and we will discuss this approach in more detail in Sections 2, 3.

aTFs can present several architectures. The relationship between the effector molecule and the aTF defines their mode of action: repression of activator aTF, activation of repressor aTF, repression of repressor aTF, or activation of activator aTF (Mannan et al., 2017). A sizable number of aTFs have been found for each category, allowing one to build biological circuits with a large variety of complex functions. However, the use of TF-based biosensors in complex applications such as the industrial scale-up of bioproduction processes or intricate biocomputing circuits has been stalled. This is mainly due to the fact that the number of metabolite-activated TFs which have been described in the literature (Koch et al., 2018) is rather small compared to the large number of compounds potentially amenable to biomanufacturing. Additionally, biosensing circuits often perform poorly due, for example, to non-specific activity, cross-talk with native biochemical reactions, leaky expression and problematic or impossible heterologous expression. It is therefore becoming increasingly clear that the number of engineered TF-biosensors and the means for their optimisation need to keep up with the growing demands of the synthetic biology community.

Here, we provide a roadmap for the design of new biosensor circuits based on aTFs that leads from gathering data and theoretical prediction to experimental validation. We also provide guidelines for the rapid prototyping of biosensor circuits with improved features using computational tools and discuss experimental validation methods best suited for this task. Additionally, we focus on two crucial applications of biosensors for current synthetic biology targets, namely production screening and dynamic regulation.

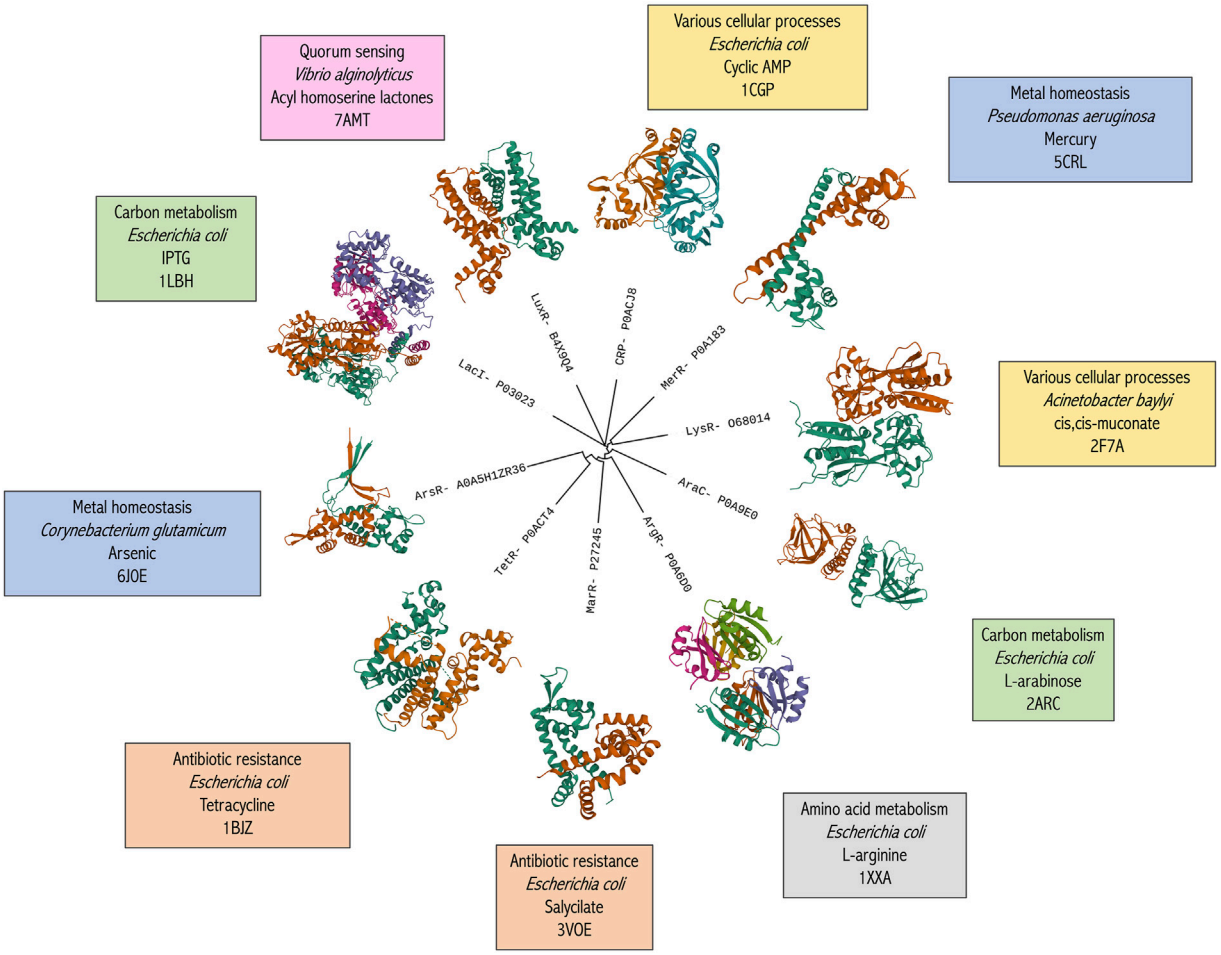


FIGURE 1
Broad phylogenetic classification of the ten most common bacterial aTF families. Family name and representative UniProt IDs are used as tree labels. Structure images were obtained using Mol* Viewer (Sehnal et al., 2021) *via* RCSB PDB. The text boxes detail, in order, the most common aTF-controlled pathway, the species name of the specific aTF, the representative effector molecule and the PDB identifier of the structure.

2 Determining the design space of detectable compounds and aTFs

2.1 Exploring the current knowledge of the biosensor space

For years, researchers have studied the regulatory networks of different cells and organisms to understand, among other things, how they react to environmental changes by controlling essential cellular activities through the expression or repression of their genes. With the advent of genetic engineering and synthetic biology, this information can nowadays be used to re-engineer and create fine-tuned genetic circuits for various purposes, notably biosensors. However, the data is often scattered and incomplete, and gathering efforts have to be made to organize and make easily available the current knowledge on the topic. Through literature and database mining, it is possible to build a dataset of transcription factors triggered by the binding of molecules as well as other types of inputs (temperature, light, pH. . .). This valuable dataset can be used to determine the known detectable input space (Koch et al., 2018), i.e., the set of molecules that can be detected by TF-based biosensors. Table 1 describes the main databases that can be used to generate this space. Note that the databases might gather different types of data: TF gene sequence, binding sites, ligands, TF regulated genes and/or structure, thus restricting data integration.

The databases in Table 1 can be considered as good starting points when compiling the list of known ligand-responsive TFs. However, some may not have been recently updated, and thus some important bits of information may be missing. In order to expand the initial set, Natural Language Processing (NLP) may be used to mine additional information from the literature helping to fill in the knowledge gap of the metabolite detectable space, i.e., the set of metabolites that can be detected through TFs-based biosensors. Bibliographic databases such as NCBI are ideal sources of input data for this type of text mining algorithm. NLP has successfully been used in other biotechnology applications, as, for example, in the context of predicting protein-protein interactions and establishing gene-disease relationships (Zeng et al., 2015), and could also be used to find aTF-ligand interactions.

2.2 Homology-based prediction to enlarge the aTFs dataset

The body of experimental data on aTFs, which is currently available in a curated form, can also be extended through

homology-based prediction of TF sequences in other species (Figure 1). Using protein sequence information of well-known transcriptional regulator families (e.g., LysR, TetR, AraC, *etc.*) (Fernandez-López et al., 2015) as a reference, annotation experiments can be made on target genomes that are publicly available (Deplancke and Gheldof, 2012). Similarly, metagenomes can be mined (Oliveira Monteiro et al., 2021) to find more information about, for example, non-culturable species. Some of the databases described in Table 1 were built, either in their entirety or in part, by using genomic or metagenomic data.

Protein structure information can also be used to discover and engineer new biosensors. However, despite the rather large number of protein structures available for certain aTFs (e.g., TetR, TrpR, AraC), many candidates still await structural characterization, a problem that negatively impacts the development of biosensors tailored to recognize arbitrary compounds of interest. Although many aTFs do show sufficient homology with existing structures and might therefore be targeted with homology modeling, the quality of these models can at times be lacking—especially in the twilight zone of sequence homology. This, in turn, limits the accuracy with which residues that are crucial for the binding of effector candidates can be predicted by docking calculations. Based on the current knowledge of aTF-ligands (Koch et al., 2018), we calculated that not more than 45% of the TFs in this dataset share more than 50% of identical residues with experimentally solved protein structures.

2.3 AI-based prediction of new aTFs

Traditional homology and structural-based methods are not the only approaches that can be used to predict novel TFs. In recent years, as in other biotechnology fields, AI-based applications have been widely adopted (Volk et al., 2020). The predictive power of AI can be exploited by combining the information available in the aforementioned databases with reference genomic databases such as NCBI, to train the algorithms for predicting new TFs. In 2021, Kim and others presented DeepTFactor (Kim et al., 2021), which was able to predict over 300 TFs in the well-studied *Escherichia coli* K-12 genome, including TFs not previously reported in databases, and which the authors were able to validate using TF and non-TF protein sequences as training data. This tool can be used to identify new TF sequences from known and new genomic and metagenomic information, in combination with sequence-based annotation tools. Other AI-based methods that identify DNA binding protein domains (Eichner et al., 2013; Mishra et al., 2019; Li et al., 2021) from sequences may also be considered for this purpose. The recent progress in the application of deep learning to the *de-novo* prediction of nucleic acid/protein complex structures, such as RoseTTAFold2NA (Baek et al., 2022) and DeepFoldRNA (Pearce et al., 2022), might make it possible to structurally validate TF/DNA complex formation and binding for those TF candidates that show no or only poor homology with known protein structures.

Protein structures open an extra layer of useful data for researchers trying to complete the atlas of aTFs. In a milestone research article, AlphaFold was published in 2021 as a Machine Learning prediction tool, able to predict protein structure using only the protein amino acid sequence as input (Jumper et al., 2021). This tool is expected to provide reliable and fast structure information that would otherwise take years to be resolved through experimental methods.

Once a putative TF has been identified, determining the TF binding site (TFBS) (i.e., the DNA region where the TF attaches to/detaches from) is the next step. Sequence homology among TFBSs of the same aTF can be used together with AI (Koo and Ploenzke, 2020; Chen et al., 2021) to predict new TFBSs. Some of the databases in Table 1 (Santos-Zavaleta et al., 2019; Castro-Mondragon et al., 2022) provide this information for each of their entries. To that end, a structural component that determines the activity of an aTF is its ligand binding domain (LBD). The practical applicability of TF-based biosensors is limited by the relatively small pool of known ligand binding domains. Designing and experimentally validating new binding domains has been a long-standing challenge. Thanks to continued work and recently introduced algorithms (Lucas and Kortemme, 2020; Polizzi and DeGrado, 2020), a viable solution seems now within reach but not yet at hand. In addition to high ligand affinity and specificity, the receptor system must be activated once bound to the effector molecule, going through some form of conformational change in order to trigger the biosensor response (Su and Hammond, 2020). This activation can be hard to engineer for new molecules and protein domains. AI might also be able to address this challenge, as it can be used not only to identify new putative TFs, but also to predict the pockets (i.e., sites) of allosteric interaction (Figure 2A) with the ligand in the LBD (Panjkovich and Daura, 2014; Greener and Sternberg, 2015; Xiao et al., 2021).

Once novel TFs and their TFBSs have been found, it is necessary to determine their most likely effector molecules. To do this, Hanko and others (Hanko et al., 2020) described an approach to identify whole metabolite-inducible systems (i.e., TF, inducible promoter and its corresponding effector). The method looks at the operon next to the predicted TF and assigns the effector molecule as the primary substrate metabolized by the operon-encoded enzymes. A structure-based solution was also made available in (Huang et al., 2018) where Huang et al. presented a computer tool that finds allosteric sites inside a structure and calculates a docking score for each molecule in a database (Figure 2B).

Allosteric sites can be good candidates for AI-mediated site-directed mutagenesis experiments (Cadet et al., 2018; Saito et al., 2018; Wu et al., 2019) (Figure 2C) trying to improve the specificity of the TF towards the ligand or adapting its affinity towards novel molecules. Machine Learning-centered structural approaches such as AlphaFold might soon lead to a breakthrough in the field of transcription factor design, in particular because the most recent iteration, AlphaFold-Multimer (Evans et al., 2021), is able to produce multimeric solutions. These can be especially relevant for the effector binding domains (EBDs) of homodimeric aTFs and are a prerequisite for predicting binding modes of effector molecules accurately. Of note are also very recent language-model-based approaches that appear to be computationally more efficient and are able to predict multimeric states even though they have not explicitly been trained on protein complexes (Lin et al., 2022).

The design process of bespoke binding domains for non-cognate ligands would also benefit from the development of faster and more reliable *in silico* docking algorithms. Recent ML-based advances, such as DiffDock (Corso et al., 2022), might provide a crucial advantage over classic methods in this context. The computational speed-up promised by these approaches might also make it possible to explicitly model water molecules and thereby further increase the predictive

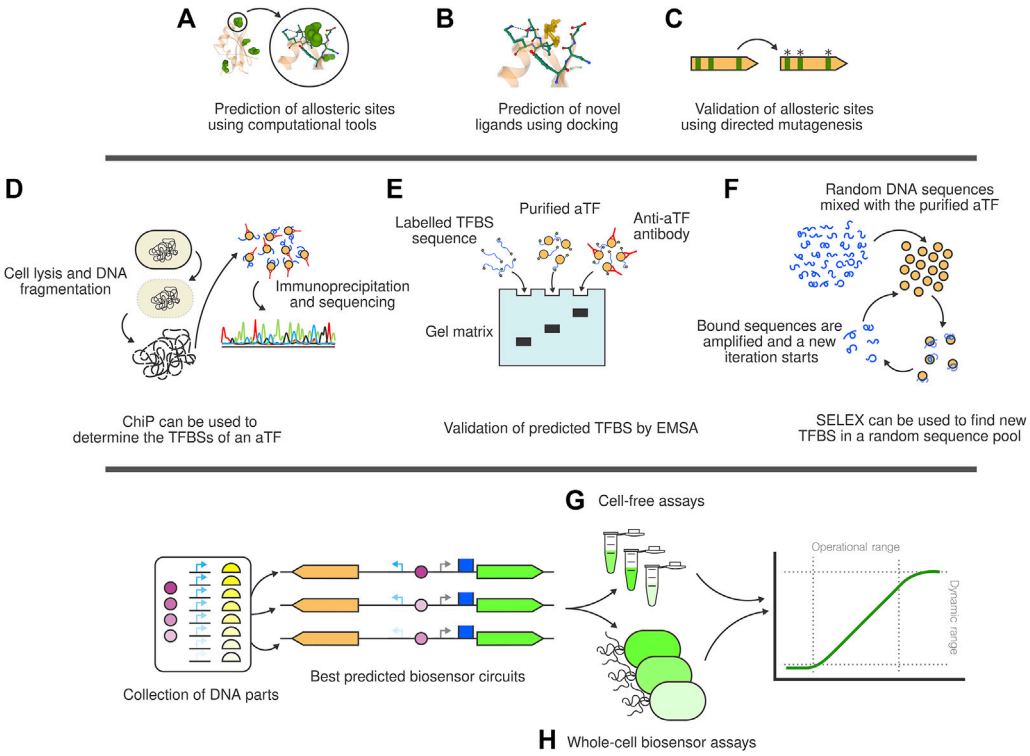


FIGURE 2
The full stack biosensor development toolbox. **(A)** Computational tools can be used to determine the allosteric pockets of interaction between ligands and aTF. **(B)** Using these pockets as reference, docking computations can be carried out to assess the affinity of the aTF towards a library of putative ligand compounds. **(C)** The allosteric site computation can be validated using directed mutagenesis to evaluate changes in affinity or specificity. **(D)** ChIP technology allows researchers to determine the TFBSs of a newly discovered aTF. **(E)** Similarly, EMSA can be used to individually validate single TFBS. **(F)** SELEX can be used to artificially obtain new TFBS to the aTF. **(G)** Cell-free assays are a quick prototyping technology to test biosensor circuits once assembled. **(H)** Whole-cell biosensor experiments allow the characterisation of the biosensor circuit in conditions closer to *in vivo* applications.

quality of the docking process, because water-mediated interactions often play a crucial role in ligand binding sites. The prediction of binding modes for possible effector molecules is necessary but not sufficient for the reliable computational design of biosensors due to the delicately tuned allosteric properties of aTFs. These can make protein engineering attempts to alter ligand-binding specificity quite challenging, given the possibility of negatively impacting signal transduction.

Unfortunately, predicting the dynamic properties of an aTF upon ligand binding and/or mutation is significantly more challenging than calculating binding modes and usually involves long time-scale molecular dynamics simulations. This problem could, in principle, be addressed with coarse-grained molecular dynamics methods that approximate fully atomistic simulations, such as the latest iteration of the Martini model, which has successfully been applied to ligand binding (Souza et al., 2020). Crucial for its applicability, however, is the availability of parameters for the ligands to be studied. For the Martini model, this issue is being addressed by the creation of a curated ligand database and the development of automated tools for the generation of coarse-grained models (Souza et al., 2020; Hilpert et al., 2022). It is, however, conceivable that the problem of predicting ligand-induced conformational changes in aTFs can be addressed with deep-learning approaches as well, which could constitute a breakthrough for the *in silico* component of biosensor-engineering.

2.4 Experimental validation of predicted biosensors

Once a putative new aTF has been identified, experimental validation is necessary to check its DNA binding site, affinity towards different ligands, and performance. For each of these tasks, some validation procedures have been defined. Some of the tools described in previous sections and in Table 1 can be used to predict TF DNA binding sites. However, for the subsequent development of a functioning biosensor circuit, it is crucial to experimentally confirm that the TF actually binds to the effector molecule and to its target DNA sequence.

For instance, chromatin immunoprecipitation (ChIP) is commonly used to assess the binding sites of the aTF anywhere in the genome, effectively assessing the regulated genes (Figure 2D). To this end, the cells are lysed, the genome is fragmented and TF-DNA complexes are isolated thanks to TF-specific antibodies (van Werven, 2006; Grainger et al., 2007). In a further step, the DNA in the complex can be sequenced to determine the TFBS.

The Electrophoretic mobility shift assay (EMSA) is a standard procedure that can be used to confirm the DNA binding site of newly discovered TFs (Alves et al., 2021; Modrzejewska et al., 2021) (Figure 2E). Basically, the purified TF is mixed with a labelled DNA probe of the putative binding sequence and ran in an agarose gel. The TF-DNA complex runs slower compared to the

free DNA sequence (Gurevich et al., 2010). EMSA, however, is not scalable. Other procedures have been used to determine, using high-throughput technologies, the DNA binding sequences of hundreds of TFs annotated in a species. Wang and others (Wang et al., 2021), for example, determined the binding specificities of 182 TFs of *Pseudomonas aeruginosa*. To do this, they purified 371 putative TFs, mixed them with randomised DNA sequences and ran high-throughput systematic evolution of ligands by exponential enrichment (HT-SELEX) (Figure 2F) for four cycles. Both EMSA and the HT-SELEX require the TF to be previously purified. This can quickly become a bottleneck. For this reason, a bacterial one-hybrid system was described in 2005 (Meng et al., 2005). In a nutshell, this procedure allows researchers to assess the binding specificities of TFs against random DNA sequences *in vivo* by linking the TF-DNA binding with the expression of positive and negative selection markers. This method does not require proteins to be purified or the availability of antibodies and provides a low-tech scalable method of finding DNA binding sequences.

2.5 Extending the biosensor space through bioretrosynthesis

In the cases where no aTFs are known to be triggered by the target compound, alternative approaches may be considered. One possibility is to enlarge the biosensor space is through retrosynthesis approaches (Delépine et al., 2016). The approach enables modulation of the specificity and dynamic range of the biosensor by introducing metabolic conversions as part of the sensing process. In this way, the number of targets that can be detected can be substantially increased since any chemical target that can be converted into a molecule for which an aTF exists becomes potentially detectable. In order to compute the metabolic pathways that can connect the target to existing aTFs, bioretrosynthesis-based approaches are used (Lin et al., 2019). Such algorithms generate a tree-like graph of biochemical conversions connecting the target to those molecules that can be detected. SensiPath (Delépine et al., 2016) is an online server that can compute the alternative extended aTF-based biosensors for any given target.

This approach has been used for instance to analyse the biosensing of the production of naringenin that is used for dynamic regulation through its conversion into kaempferol (Boada et al., 2020). The study showed how the dynamic range of the resulting biosensor could be adjusted up to industrial levels of 1 g/L in a bioreactor, in a way that would have been more challenging by using direct biosensing of naringenin.

The approach has also been used systematically in order to develop a protocol for the development of cell-free biosensors through metabolic and genetic layers (Soudier et al., 2022). The authors proposed a standard methodology based on computer-aided design (CAD) that combined the design of a perceptron-like genetic device (Pandi et al., 2019a) with the automated selection of enzymes for the metabolic pathway through the Selenzyme algorithm, an online tool that suggests best candidate enzyme sequences based on the biochemical conversions in the pathway (Carbonell et al., 2018a).

Other engineering efforts have been showcased which focus on different areas. Biocomputing, for instance, would also benefit from a wider range of TF-based biosensors. To that end, Rondon and others (Rondon et al., 2019) created 27 new synthetic TFs starting from 6 core TF domains, 7 DNA recognition domains and 7 operator regions that are able to detect 5 different ligands.

3 Biosensor design

3.1 Biosensor library characterization and fine-tuning

The predictions and the *in vitro* and *in vivo* validation experiments provide essential information to biosensor designers. An intermediate step between prediction and experimental validation and actual whole-cell biosensor construction can be achieved by cell-free assays. The technique relies on cell extracts containing all the necessary cellular components for protein expression. It removes cellular maintenance, growth, other native processes and all cellular unknowns from the equation, allowing the experiment to be essentially focused on the synthetic circuit and behavior designed by the researchers (Hodgman and Jewett, 2012). Cell-free assays (Figure 2G) provide a simple and standardizable approach to quickly test biosensor genetic circuits (aTF, aTF promoter, reporter, operator. . .) *in vitro*. Examples of successful cell-free biosensors have been described, that detect chemicals such as quorum sensing molecules (Wen et al., 2017), rare sugars such as D-psicose (Pandi et al., 2019b) and water contaminants (Jung et al., 2020), among others.

However, building a biosensor circuit that works inside the cell (Figure 2H) and that provides a measurable output is the litmus test of the successful functionality of the new aTF. In its most basic form, a biosensor circuit will express the aTF and a reporter (e.g., a fluorescent protein). The latter is under the control of a promoter containing the DNA-binding sequence recognized by the TF. The binding of the effector molecule with the aTF regulates gene expression of the reporter gene. Some considerations to bear in mind during the design of such metabolite-responsive biosensors, specifically for the aTFs, can be found in (Liu et al., 2017).

Once a circuit has been designed and built, the next step is to determine how well it works under a range of conditions; that is, to characterise the biosensor. Characterisation experiments determine, among other parameters, the dynamic range, the operational range and the dynamic response of the biosensor (Mannan et al., 2017). It is likely, however, that the initial biosensor design is not fit for its intended purpose and tweaks need to be done in the circuit for the biosensor's performance to meet the designer's criteria (e.g., dynamic range, operational range, specificity, speed of response. . .). The dynamic and operational range of the biosensor can be engineered by changing the expression levels of both the TF and the reporter gene through RBS (De Paepe et al., 2018) or promoter (Sonntag et al., 2020) engineering. The number and position of the aTF operator region(s) can also be used to modulate the dynamic range of the biosensor (Xu et al., 2020). On the other hand, the generation of chimeric aTFs, obtained by merging the DNA binding domains and ligand binding domains from different genetic sources has also been successfully tested to modulate the specificity towards other ligands (De Paepe et al., 2019). A similar approach based on the high-throughput fusion of periplasmic binding proteins and DNA binding domains was presented in Juárez et al. (2018). A complete overview of the different approaches that can be used to fine-tune the initial biosensor was provided by De Paepe et al. (2017).

Apart from poor inherent behaviour of the aTF, the compatibility of the heterologous host with the transcription factor is often the main culprit when the performance of the newly generated biosensor is inadequate. The portability of genetic circuits between species has been challenging synthetic biology since its early stages. In bioproduction projects, it may be possible that the best host for the production of the target compound is not suitable for the expression of the biosensor circuit. In the native species, transcriptional regulation often implies a complex regulatory system composed of several membrane transporters, inhibitors, activators and cofactors (Carpenter et al., 2018). For well characterised regulators, these may be known but newly found aTFs may have unknown necessary components. The challenge is bigger when the aTF system is transferred from phylogenetically distant species (e.g., from plants to bacteria). Steps that can be taken to improve the performance of biosensor circuits include: optimization of gene expression, selection of an adequate reporter system (see next Section) and the incorporation of additional genetic modules which may entail the addition of compound importers, exporters, leak dampeners and other types of signal modulation devices (e.g. amplifiers, inverters…) (Miller et al., 2022).

3.2 Biosensor reporter selection

An important and often overlooked factor in designing and building new biosensor circuits is the selection of the reporter gene. Fluorescent reporters are the most common choice. Classic fluorescent proteins (FPs) such as GFP and RFP are simple to assemble in genetic circuits, easy to measure and do not rely on any metabolic substrate (other than oxygen) to work. Most of the experimental references detailed in this work use FPs as reporter systems and this type of biosensor is predominantly used in the field. Different FPs offer different characteristics to the biosensor designer. One should consider, among other factors, the excitation and emission wavelength, the maturation time (Shaner et al., 2005) and the half-life of the matured protein [as the FP can sometimes be too stable and rendered useless in real-time applications such as biosensing (Andersen et al., 1998)]. Bioluminescent proteins are other alternative reporters (Nourmohammadi et al., 2020; Hansen et al., 2021) that rely on biochemical reactions emitting photons as products. Compared to fluorescent proteins, bioluminescent reporters do not rely on the measuring equipment to excite a fluorophore, which leads to less background emission and higher sensitivity.

Nevertheless, there are other alternative reporters that offer different features that may be more appropriate for specific purposes. Before fluorescent and luminescent signal detection equipment became ubiquitous in molecular biology laboratories, colorimetry was often the most efficient way to detect biological processes using the naked eye or simple absorbance measurements. Biosensors have been developed using the colorimetric reporters lacZ (Choi et al., 2013; Li et al., 2017; Hansen et al., 2021) and the carotenoid pathway (Yoshida et al., 2008; Watstein et al., 2015). In recent years, the violacein pathway has gained popularity as a tunable route where each intermediate compound can act as a measurable reporter (Watstein et al., 2015; Hui et al., 2020; Guo et al., 2021). Reporters that utilize electrical signals have also been proposed and are derived from electrogenic bacteria (Golitsch et al., 2013; Webster et al., 2014; Zhou et al., 2017) or are obtained *via* a synthetic electron transport chain (Atkinson et al., 2022), where the generation of an electric current allows for faster responses and actuation than can be achieved with protein expression-based systems.

A thorough consideration of the features and issues of each reporter category should be taken into account before committing to a reporter. A comparison of 8 different reporters of three categories can be found in (Lopreside et al., 2019). In short, enzymatic reporters (LacZ and bioluminescent reporters) can have the fastest response and the lowest detection limit for the target metabolite which is perfect for biosensors requiring precise and quick measurements. However, these advantages require the cells to be lysed and the enzymatic substrate to be added to trigger the reporter reaction which does not allow the users to perform continuous quantification experiments. On the other hand, fluorescent proteins can present high media- and cellular autofluorescence and a slower response. Nevertheless, FPs have managed to become the first choice of many researchers thanks to, among other things, the possibility to simultaneously use multiple reporters with different emission patterns (green, red, blue…), the ease of use and the stability of the proteins.

3.3 Directed evolution of aTFs

Using smart design and predictions can be a good way to consolidate the design of the biosensor. However, when trying to, for instance, increase the sensitivity of the aTF towards new ligands, it may be necessary to apply directed evolution techniques on the gene sequence. An example can be found in the work of Rottinghaus and others, where new variants were engineered from promiscuous amino acid-specific TFs to specifically detect similar amino acids and neurotransmitters (Rottinghaus et al., 2022). A review on the evolvability of TF-based biosensors can be found in (Umeno et al., 2021). This may produce several hundreds of variants, and most of them may show equal or poorer performance, which requires a faster and simpler approach that quickly identifies and isolates good performers and discards variants that do not meet the designer's criteria. Several methods have been described for this task.

Fluorescence-activated cell sorting (FACS) isolates cells based on fluorescence emission detected by a flow cytometer. This technique allows researchers to select, from a cell population, the individuals that perform best. FACS has previously been used to reduce the affinity space of an aTF library, thereby creating a sensor for L-histidine and L-arginine which is unable to detect L-lysine (Della Corte et al., 2020). Machado and others changed the specificity of a protocatechuic acid (PCA) biosensor to instead detect vanillin and 3,4-dihydroxybenzaldehyde using this technique (Machado et al., 2019).

The promoter under control of the effector molecule can be put in front of other types of markers. In Collins et al. (2006), the researchers managed after several rounds of directed evolution with error-prone PCR, to select variants of the aTF LuxR by isolating cells that survive on chloramphenicol and discarding variants that survive on carbenicillin by controlling the expression of cat gene and the bli gene (β-lactamase inhibitory protein, inhibits the *bla* gene) using the aTF binding site. The new variant managed to respond to new ligands, but no longer responded to the original effector.

Less common techniques have also been used to evolve aTFs to recognise new ligands. One example is a method called compartmentalised partnered replication (CPR) (Ellefson et al., 2014). In 2018, researchers generated a synthetic phylogeny from the aTF TrpR using this approach (Ellefson et al., 2018). Very briefly, CPR can be used to evolve DNA parts *via* coupling the allosteric effector activity to the expression of Taq polymerase *in vivo*. The

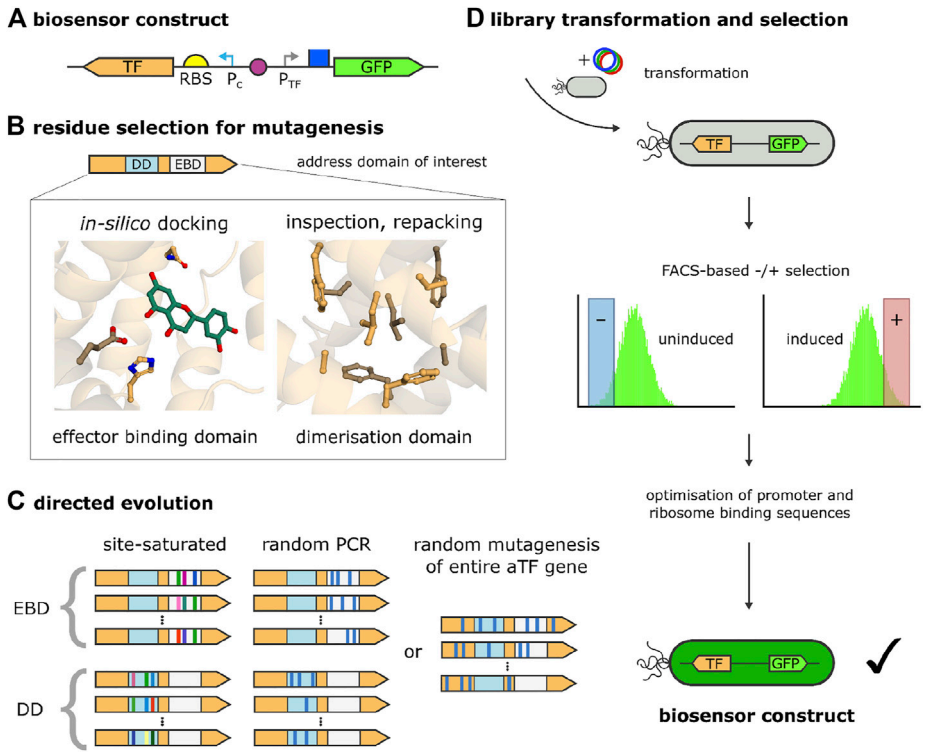


FIGURE 3
Biosensor development by directed evolution of the aTF. **(A)** Biosensor construct on one plasmid, composed of the sensor component, which, in turn, contains the transcription factor (TF) gene with a ribosome binding site (RBS) and a consecutively active promoter (P_C), and the reporter component, comprised of a TF-inducible promoter (P_{TF}) and a reporter gene (green fluorescent protein, GFP). **(B)** Biosensor mutagenesis strategies. Depending on the availability of structural information, a domain of interest of the TF (e.g., the effector binding domain, EBD, or the dimerisation domain, DD) are inspected or modeled and residues or sequence stretches for mutagenesis are selected. To guide mutagenesis, ligands can be placed into the EBD with *in silico*-docking and side chains can be designed and repacked (e.g., in the DD). **(C)** Directed evolution strategies. Based on the residues or sequence stretches selected, the domain of interest is addressed (EBD or DD in the current examples) by rational approaches, i.e., site-directed mutagenesis or random/error-prone PCR. Alternatively, the complete aTF gene is targeted with random mutagenesis in the absence of structural information, or if non-intuitive effects should be probed (e.g., those affecting allostery or DNA binding). During site-directed mutagenesis, combinations of amino acids are generated for selected, fixed sequence positions (this is represented in the picture by different colors at identical sequence positions). **(D)** The library of sensor constructs obtained by directed evolution is transformed into cells, which are then subjected to (usually) multiple rounds of negative selection (non-induced sensor in the absence of effector) and positive selection (induced sensor in the presence of effector), facilitated by fluorescence-activated cell sorting (FACS) [after (Machado and Dixon, 2022)]. The best candidates are further optimised with respect to other genetic elements determining the fidelity of the bio-sensor, such as RBS and promoter sequences.

variants that manage to produce larger concentrations of Taq will then be amplified using emulsion PCR.

Apart from the methods described so far, also phages have successfully been applied to evolve TFs. Phage-assisted continuous evolution (PACE), described in 2011 (Esvelt et al., 2011), can be used to continuously evolve any gene that can be coupled to pIII production in *E. coli*. pIII protein is required for phage infection. After mutagenesis, only the gene variants able to induce enough pIII production will propagate and enter the next cycle. A modification of this system was applied to the evolution of a TF (Brödel et al., 2020) that regulated the expression of another gene essential to phage propagation.

3.4 Automating the design of new biosensors

As we have described, building a new biosensor circuit using a new aTF requires many experimental steps from ligand and DNA affinity validation, to parametric characterisation, to fine tuning inside the heterologous host. All these steps can quickly become a bottleneck, especially if several ligands and aTFs are tested at the same time for different purposes. For this reason, high-throughput automated construction must be considered by the biosensor circuit designer. An application is showcased in (Tenhaef et al., 2021) where researchers rationally designed and characterised a library of aTF biosensors based on the Lrp regulator. The parts (promoters, RBS...) can also be engineered using automation, generating thousands of variants and combinations (Hossain et al., 2020). To that end, rapid prototyping strategies through automated Design-Buid-Test-Learn pipelines (Carbonell et al., 2018b) such as those implemented in biofoundry facilities (Tellechea-Luzardo et al., 2022) should alleviate current bottlenecks in the design of new biosensor circuits.

The discussion above focuses on general principles related to biosensor design. In the following section we will explore instructive examples from the literature that highlight practical aspects and possible limitations in current biosensor projects.

3.5 Three examples of biosensor design and optimization

Exemplarily, we will discuss here three biosensors that were successfully developed by employing current experimental and computational approaches and which also underline the relevance of structural insights into the aTF.

3.5.1 Directed evolution strategy for biosensor engineering

As alluded to above, in order to alter the properties of an aTF (e.g., its ligand specificity, dynamic range, *etc.*) directed-evolution approaches have proved to be indispensable (see Figure 3 for a general overview). Directed evolution can include rational approaches which, informed by structural insight into the effector binding domain (or any other domain of interest), apply error-prone PCR within the whole region, or saturation mutagenesis to residues identified as crucial for ligand binding or other properties (Figures 3A–C). In the absence of structural or homology information, the entire protein can be subjected to random mutagenesis as well. This, however, leads to much larger strain libraries that are still unable to sample efficiently the larger sequence space, despite the higher experimental burden. An advantage of this approach is that it can reveal beneficial mutations that are less obvious than those directly affecting ligand binding, and might, among other effects, induce subtle changes in the allostery of the TF. In general, however, it is prudent to limit the size of the search space with the help of structural information, not least of all because the development of a complete bio-circuit requires the optimization of additional genetic elements, which further increases the combinatorial complexity of the design problem. This, for example, includes selecting RBS and operator sequences that perfectly harmonize with the engineered aTF.

3.5.2 Development of biosensors for polyphenols

Machado et al. (2019) describe the development of a biosensor for detecting protocatechuic acid (3,4-Dihydroxybenzoic acid, PCA) by using the aTF PcaV from *Streptomyces coelicolor*. Their setup uses a two-plasmid approach, with which the expression of a GFP reporter gene is under the control of the (constitutively) expressed PcaV repressor protein. In its initial form, the biosensor was responsive to a narrow range of hydroxyl-substituted benzoate derivatives and displayed only a modest dynamic range for these effector molecules, which is commonly observed with aTFs. By applying directed evolution, Machado et al. (2019) were able to change the selectivity of PcaV so that it was able to recognize the phenolic aldehyde vanillin (4-Hydroxy-3-methoxybenzaldehyde) instead of its cognate effectors, which is quite remarkable, given the high chemical similarity between this compound and the original group of ligands. Crucial to this approach was the availability of crystal structures for PcaV, which permitted the authors to restrict the directed-evolution procedure to seven side chains that, in the PcaV protein, are in close contact with the effector PCA. To screen the strain libraries for functional mutants, the researchers developed a fluorescence-activated-cell-sorting (FACS) counter-selection protocol. This applied several rounds of negative selection (for variants that remain uninduced in the absence of the effector) and positive selection (for candidates that are also strongly inducible by the proper ligand) (see Figure 3D). Further analysis of the most promising variants revealed that a total of only three mutations sufficed to change effector recognition from PCA to vanillin and related aromatic aldehydes.

This example, in which the specificity of an aTF was successfully engineered, demonstrates the effectiveness of structure-based, directed-evolution approaches. But, as noted above, the properties of a complete aTF-based biosensor are affected by multiple variables, such as the sequences of operator and ribosome binding sites, which can turn the optimization of the complete system into a daunting task. This particular problem was addressed by Berepiki et al.(2020) while revisiting the PcaV system just described. To this end, they merged the constitutively active pcaV gene with a GFP-reporter-gene construct on a single plasmid in order to facilitate the optimization process, which involved the systematic variation of the genetic components of the sensor, i.e., the constitutively active promoter for the expression of PcaV, the repressible PcaV-regulated promoter of the reporter gene (GFP), and the ribosome binding site for the translation of the GFP transcript. By randomizing selected positions in these genetic elements, the authors obtained three distinct sequence libraries, whose simultaneous exploration and optimization would have resulted in more than ten thousand combinations to test. In order to tame this combinatorial complexity, the researchers applied a design-of-experiments (DoE) approach which determines the ideal combination of experiments that most efficiently probe the possibility space while reducing the experimental effort. With this strategy, the authors were able to reduce the experimental setup to just thirteen distinct combinations of the genetic elements.

With the DoE approach employed, the authors assessed the impact of each variable (constitutive promoter strength of the aTF; ribosome binding site of the reporter gene; and promoter strength of the reporter gene) in a semi-quantitative way *via* statistical modeling, which even revealed non-linear effects among the parameters, and derived general design rules for repression-based aTF biosensor systems. According to the authors, the first step should be to identify the strongest combination of promoter-operator and RBS; Then, the regulator (aTF) expression should be fine-tuned by testing a wide range of expression levels. As a last step, if the dynamic range of the sensor is still not satisfactory, the RBS, which drives signal output, should be weakened.

The authors further demonstrated the generalizability of the DoE approach by optimizing a biosensor for ferulic acid that they had previously developed. With just twelve experiments, they were able to improve the maximal output signal of this biosensor by a factor of 32 and the dynamic range by a factor of five.

3.5.3 Biosensor development for monitoring L-cysteine levels in cells

As discussed, structural or homology information about the aTF can be crucial for limiting the experimental effort in biosensor-design, which is also highlighted by the recent work of Gao et al.(2022). They present the development of a novel biosensor for L-cysteine, starting from the L-cysteine-responsive transcriptional regulator CcdR. With this sensor, the authors wanted to efficiently detect and analyze cysteine-overproducing strains obtained from rational or random mutagenesis libraries. One bottleneck in directed-evolution strategies is the screening stage, for which the authors devised a high-throughput screening (HTS) method. Like the PCA biosensor discussed above, Gao et al.(2022) developed a construct that constitutively expresses the aTF (CcdR), which, in turn, acts upon a GFP reporter gene under the control of the CcdR-specific operator ccdA. The initial construct showed proper dose-response behavior and specificity with respect to the effector L-cysteine, but low sensitivity

and a narrow dynamic range, which rendered it unsuitable for the screening of large microbial variant libraries. In order to improve the initial design, the authors optimized the genetic components of the sensor system by directed evolution of CcdR, and by combinatorially optimizing promoter and RBS sequences for its expression.

For improving the properties of CcdR, the authors applied a semi-rational design strategy, made possible by the high homology of CcdR with other, structurally resolved, members of the FFRP-like (feast/famine regulatory protein) aTF family.

Because cysteine is already the cognate ligand for CcdR, the authors focused not on the effector binding but rather on the dimerization domain of the aTF, arguing that improving the dimerization properties of the repressor would increase its biological activity. Based on homology, the researchers selected eight sites in the putative dimerization interface of CcdR for saturation mutagenesis, followed by FACS-based selection, similar to the strategy described above for the PcaV-based sensor. This approach eventually led to the identification of a single-point mutant with significantly increased responsiveness to L-cysteine and a higher signal-to-noise ratio than the wild type protein. With the help of an AlphaFold-generated model (Evans et al., 2021), the authors suggest that these findings are due to improved hydrophobic interactions in the dimerization interface of the mutant.

Next, the authors tested the ideal combination from two promoter and two RBS sequence variants in order to further improve the switching dynamics and the sensitivity of the biosensor, and thus faced much less combinatorial complexity than was encountered with the approach discussed above for the PCA biosensor.

In order to facilitate the screening for L-cysteine-overproducing strains in directed evolution experiments, Gao et al.(2022) incorporated their novel L-cysteine biosensor in an HTS platform that includes the transformation of mutagenesis libraries in cells containing the biosensor, followed by FACS-screening, colony plate screening, microplate reader analysis and finally fermentation of promising candidate clones. The effectiveness of the biosensor was demonstrated by utilizing this HTS platform in the directed evolution of L-serine-acetyltransferase, the key enzyme in the biosynthesis of L-cysteine which catalyzes its rate-determining step. By using this approach, the authors were able to directly correlate enzymatic activity with the L-cysteine levels in the cells, thereby circumventing the normal time-consuming, low-throughput sorting process. Starting from a mutant CysE library that they produced *via* error-prone PCR, they succeeded in identifying a CysE double mutant with a 7-fold increased activity, and a single mutant, whose L-cysteine producing capability was 2.7-fold higher than wild-type levels.

After demonstrating that the sensor system can in principle be applied for the screening of strains, the authors assessed its ability to handle large mutant libraries. To this end, they subjected the biosensor-containing strain to ARTP (atmospheric and room temperature plasma) mutagenesis and applied the HTS protocol, which reduced an initial amount of ten million cells to ten strains that all featured higher production levels for L-cysteine than the control.

3.5.4 Development of biosensors for benzylisoquinoline alkaloids: Improving negative selection with the SELIS procedure

The previous examples relied on FACS for the selection of biosensor candidates. However, screening for strongly repressing variants by negative selection can be challenging. d'Oelsnitz et al. (2022) describe how this step can be significantly improved by suppressing the presence of dead or inactive cells, which can corrupt the cell sorting procedure. This is accomplished by making cell survival directly dependent on the repression activity of the biosensor with a method they call SELIS (seamless enrichment of ligand-inducible sensors). The approach also enables counter-selection against variants that are activated by non-target ligands, i.e., screening for repressor specificity and selectivity.

The idea behind SELIS is to add to the regular construct—consisting of the consecutive expression system for the repressor and the repressor-regulated reporter gene (see Figure 3)—One additional regulatory circuit that prevents cellular growth if the repressor only incompletely prevents gene expression in the absence of the effector. This is accomplished by growing cells in the presence of zeocin during the negative selection step, while resistance against zeocin is provided by *Sh Ble* in a way that depends on the full repression capabilities of the biosensor-candidate. To this end, the expression of *Sh Ble* is repressed by λ cI, which, in turn, is regulated by the same repressor-sensitive operator sequence as the reporter gene. If the repressor is fully active in the absence of an effector (or in the presence of a non-target ligand), λ cI cannot be formed and *Sh Ble* will be expressed, which ensures zeocin resistance and cell survival. Surviving cells can then be positively selected by plating them on zeocin-free agar plates in the presence of the effector and screening for strongly fluorescent colonies.

The authors demonstrated the effectiveness of their SELIS approach by designing sensitive and selective biosensors for a group of pharmacologically relevant benzylisoquinoline alkaloids (BIAs), namely tetrahydropapaverine, papaverine, glaucine, rotundine, and noscapine. As a starting point for biosensor development, the approach focussed on multidrug-resistance regulators that control the expression of multidrug-efflux pumps, specifically RamR from *S. typhimurium*.

The authors were able to devise an efficient mutagenesis strategy due to the availability of a crystal structure of RamR bound to the alkaloid berberine, which is structurally similar to the BIAs selected for the study. Based on this structural information, they targeted five helices surrounding the effector-binding region by creating five distinct libraries. In each library, three residues were chosen for site-saturated mutagenesis. In independent experiments, the authors also applied error-prone mutagenesis to the entire RamR gene, resulting in libraries with two mutations per gene, on average.

Starting with wild-type RamR, which displays an inherently high promiscuity for structurally diverse compounds, the authors were able with just four rounds of directed evolution, to produce highly specific biosensors that each showed >100-fold preference in binding for their cognate ligands. At the same time, each biosensor also displayed high sensitivity (<30 mM) for its target compound.

Having demonstrated the advantages of the SELIS procedure over the more traditional and purely fluorescence-based selection approach to biosensor engineering, the researchers showed how their newly evolved biosensors can be applied to the engineering of metabolic pathways. They chose the biosynthesis of tetrahydropapaverine (THP) as an example, which, in plants, involves a complicated multistep process catalysed by an oxidase and four O-methyltransferases.

Using the previously evolved THP-specific biosensor to screen for THP-producing strain variants, the authors aimed to evolve a methyltransferase from *Glaucium flavum* (GfOMT1) into an enzyme that is capable of methylating all four phenolic hydroxy

groups of the substrate norlaudanosoline (NOR) in a single step to directly yield THP, thereby circumventing its complex biosynthesis and allowing for the efficient production of this pharmacologically relevant compound. To this end, they used error-prone PCR to generate mutagenesis libraries of the GfOMT1 gene, which resulted in enzyme variants with two mutations on average. After cotransforming plasmids with the THP-specific biosensor and GfOMT1 into *E. coli* cells, the researchers then selected strains based on high fluorescence in the presence of the substrate NOR, indicating the production of THP. Strikingly, after only three rounds of directed evolution, variants could be identified that completely converted the precursor NOR to THP, thus demonstrating the effectiveness of the biosensor for metabolic engineering projects.

3.5.5 Prerequisites for the successful directed evolution of aTFs

These examples highlight the significant progress that the combination of directed evolution and high-throughput fluorescence-based selection methods have created in the development of biosensors, but also show, as a recurring theme, the importance of reliable structural information about the aTF needed to accelerate the development process or even render the experimental effort feasible. This is especially true if an aTF needs to be engineered to recognize a non-cognate ligand for which there is no natural counterpart.

For the foreseeable future, experimental structure elucidation will progress at a much slower pace than the discovery of new sequence information, which emphasizes the need for structure-prediction methods that are reliable even in the absence of sufficient homology with solved structures.

Ideally, these efforts will lead to an automated protocol that hides the details of the modeling process from the synthetic biologist and automatically proposes mutagenesis libraries for a given objective.

4 Improving bioproduction using biosensors

aTF-based biosensors have a range of applications (e.g. diagnostics, environmental pollutant detection, biomaterials, health wearables. . .) (Moraskie et al., 2021). Arguably one of the most important applications is the use of biosensors in bioproduction. The world's current production of chemicals is expected to double by 2030 (Nijman and Halpaap, 2019). Nevertheless, the production of materials, fuels, pharmaceuticals, fertilisers, foods, and other types of chemicals is still heavily reliant on traditional chemical synthesis based on unrenewable, polluting, fossil fuels (Naidu et al., 2021). Using microorganisms to substitute the chemical synthesis and move to greener, bio-based processes is a thriving field known as bioproduction (Zhang et al., 2017). The goal here is to improve the efficiency with which products are produced in comparison to traditional chemical approaches and to provide pathways for the production of novel compounds for which no synthetic routes exist. Bioproduction of chemicals and materials provides a renewable and economically viable alternative, easing the transition towards a circular "bioeconomy" (Cann, 2016). Next we detail two applications of biosensors that aim to improve bioproduction: screening and dynamic regulation.

4.1 aTF-based biosensors for screening

One challenge preventing the broader adoption of current bioproduction strains is that they often perform poorly (low yield, slow production rate. . .) in industrial settings. Screening through hundreds of constructs and selecting the best performers is one of the most important bottlenecks. Thanks to advances in genetic engineering and automation, researchers can nowadays create hundreds of production strain variants in a short period of time. However, the next step involves screening those strains for production and yield. This can become a barrier since traditional methods like HPLC or LC-MS do not scale up easily. Biosensors can provide a simple, fast and affordable solution to the screening process of production strains by linking the biosensor output to the synthesis of the desired molecule (Kaczmarek and Prather, 2021). Biosensor-based screening can be used to screen for improved enzyme performance in a newly discovered or an engineered enzyme pool (Figure 4A) or for the screening of engineered genetic circuits of those enzymes and DNA parts (Figure 4B).

In order to identify biosensors with the appropriate sensitivity and dynamic range, several screening assays based on biosensors have been described. One of the simpler approaches involves the use of well plates and fluorescence detection to determine the production level of each variant (Yang et al., 2018; Zheng et al., 2018). Automation can increase the throughput of this assay by handling the library preparation, transformation and fluorescence measurement in each of the wells of the assay. Other approaches based on more advanced equipment [e.g. FACS (Liu et al., 2015), droplet-based screening (Siedler et al., 2017)] have also been tested. Tuning the dynamic range of biosensors *via* the modification of the promoter driving the expression of the reporter gene regulated by the aTF has also been proven possible (Chen et al., 2018). Biosensor mediated screening has also been described for large, multi-level CRISPRi experiments to fast-track genomic-level down-regulations that redirect the carbon flux towards the target metabolic route (Wang et al., 2023).

4.2 aTF-based biosensors for dynamic regulation

Engineered microorganisms often adapt poorly outside laboratory conditions, due to external factors originating from the conditions of production (e.g., large-scale bioreactors), such as pressure, acidity changes, accumulation of toxic metabolites, agitation, nutrient availability and heat transfer, among others (Wehrs et al., 2019). The strains are often unresponsive to external stimuli that may be present during industrial fermentation. This stress can cause undesired effects; for instance, the strain might stop producing the target compound by mutating or expelling the heterologous pathway (Wu et al., 2016). These adverse effects will lead to production processes that perform poorly when tested during the scale-up phase, thereby preventing the translation of many bioproduction projects into economically feasible industrial processes (Hartline et al., 2021).

Different approaches have been proposed to overcome those issues and increase the viability of large-scale cell-factory projects. Dynamic regulation is one of these strategies for controlling the production of key molecules, often found in nature and finely optimised through evolution. Similarly, biosensor-based dynamic

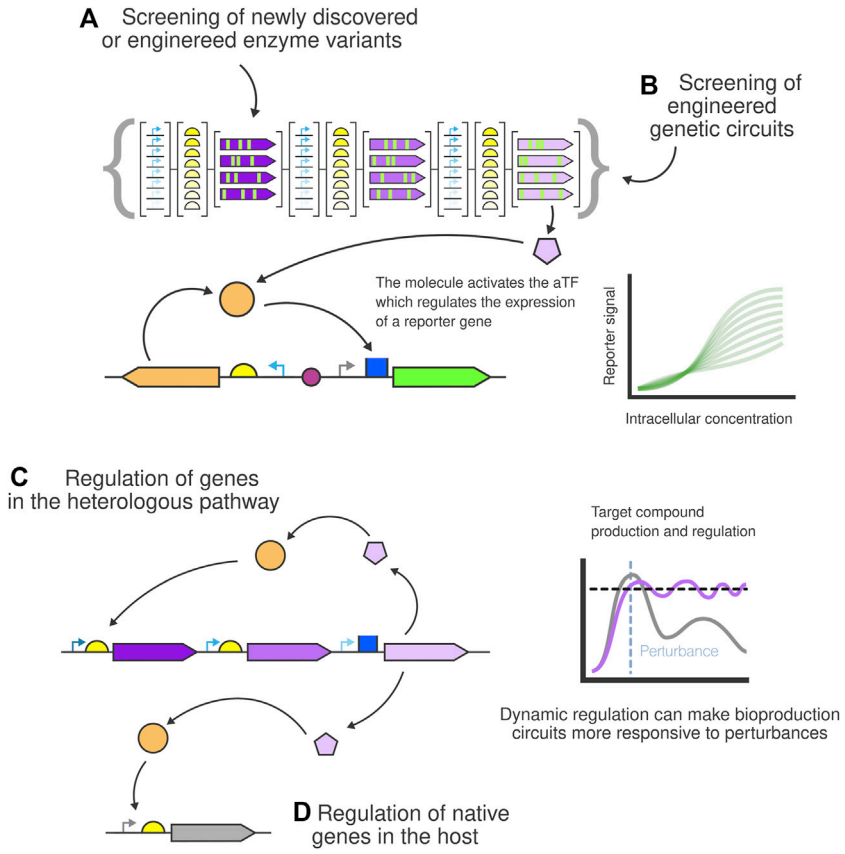


FIGURE 4
Bioproduction biosensor-based screening and dynamic regulation for bioproduction. New **(A)** enzyme variants and **(B)** genetic circuits can easily be screened for production by linking the production levels of the target metabolite to the output of the biosensor reporter. Dynamically regulated strains can be built using biosensors to control the production of the compound by regulating **(C)** the metabolic pathway and **(D)** native genes of the host.

regulation of microbial production pathways is a strategy that can be used to control genetic circuits based on a feedback loop that regulates the production of a target metabolite to keep its concentration at desired levels (Teng et al., 2022). The biosensor detects the presence of the metabolite and triggers the activation or inhibition of certain genes in the metabolic pathway (Figure 4C), making the system more responsive to possible detrimental conditions (Stevens and Carothers, 2015; Liu and Zhang, 2018; Hartline et al., 2021). To obtain the level of precise regulation required for some applications, simple single TF biosensors may not be enough. In this case, more complex circuits can be built, adding extra layers of complexity. For example, a plausible iteration would be to add another input molecule needed to trigger the desired reaction. This type of mechanism has been tested successfully with standard genetic regulator parts and inducers (Bordoy et al., 2019). Much more complex architectures are possible, mixing other common DNA parts (Nielsen et al., 2016). These approaches can be combined with native chromosomal gene regulation through direct control (Figure 4D) or by using CRISPRi (Huang et al., 2016; Wu et al., 2020) or antisense RNAs (Kim and Cha, 2003; Yang et al., 2018) that bind and repress native genetic pathways.

Growth-coupled production ties the target molecule (or one of its intermediate compounds) to the production/consumption of some essential cellular metabolite (Orsi et al., 2021). An example can be found in Wang et al. (2019) where the biosynthesis of pyruvate was limited to the heterologous pathway introduced into the cells. To accomplish this, the researchers deleted the native biosynthesis routes known to produce pyruvate. This essentially means that pyruvate will only be available as a byproduct of the heterologous pathway producing the target compound. In another example, Zhou and others (Zhou et al., 2021) built a 3-layer system that produces (2S)-naringenin, by regulating the essential compound malonyl-CoA. They used the TF FdeR, which is activated by the presence of (2S)-naringenin, and PadR TF, which is activated by *p*-coumaric acid, as feedback regulators of the (2S)-naringenin pathway. Initial low concentrations of (2S)-naringenin allow malonyl-CoA to be used in fatty acid (FA) biosynthesis pathways, favouring cell growth. Higher concentrations of naringenin, represses these FA synthesis routes, slowing cell growth and increasing the availability of malonyl-CoA for the production of more (2S)-naringenin. Through several rounds of optimization, including directed evolution of biosensors and optimization of fermentation conditions, titers above 500 mg/L from glucose were obtained in 5-L bioreactors using an *E. coli* chassis. Other strategies decouple the production phase from the growth phase of the microbial culture. For example, to improve the production of glucaric acid (GA), the glycolysis pathway can be repressed using the accumulation of *N*-Acyl homoserine lactone (AHL) in *E. coli* (Doong et al., 2018). Pyruvate-responsive circuits

were also built in *Bacillus subtilis* for the regulation of glucaric acid production (Xu et al., 2020). Pyruvate induces the expression of the GA pathway and suppresses glycolysis. Finally, malonyl-CoA, another essential cellular metabolite, can be used as the key regulator. In (Xu et al., 2014), Xu et al. built a regulation circuit that activates the synthesis of malonyl-CoA and represses the fatty acid synthesis pathway that consumes it when the compound is in low concentrations and *vice versa*.

Even though the aforementioned examples show that successful dynamic regulation is possible, true industrial utilisation of dynamic regulation circuits and scale-up is still rare in bioproduction projects. Tested under homogeneous laboratory conditions, dynamic regulation bioproduction circuits are not normally designed to react to environmental changes and, therefore, the resulting strains adapt poorly to the different conditions during the scale-up process (Neubauer and Junne, 2010). Among other reasons we can include the lack of functional and well-characterised biosensor circuits against a wider range of target molecules and the inherent difficulty of dynamically controlling complex cellular, enzymatic, genetic and molecular networks which can show unwanted crosstalk between different components. In order to address these shortcomings, a different type of regulation relies on automatically reacting to these conditions and controlling the production accordingly (Boada et al., 2020).

5 Conclusions and future perspectives

In this work we reviewed recent advances in the prediction, design and validation of biosensor circuits focused on aTFs and whole-cell implementations. This type of biosensor can be used to detect a wide range of molecules (e.g., ions, sugars, drugs, hormones. . .) for different uses, including environmental, medical and industrial applications. We showcased the role of biosensors in the development of new-generation bioproduction devices, driving forward important steps for bioproduction advancement such as screening and dynamic regulation of producer strains. aTF biosensors, however, suffer from certain shortcomings that will need addressing in future research endeavours. To that end, we detail some of the possible routes towards better biosensors.

As described in this article, several methodologies can be used to reveal as-yet-undiscovered aTFs that may behave closely to what a biosensor designer might expect. However, associated tools and databases only offer fragmented information and it is up to the designer to gather and trust the results from different sources. Therefore, a single computational tool that takes this burden from the researcher and provides confidence metrics is needed. Even though directed mutagenesis has shown great potential to improve the performance of a given aTF, it is still not possible to accurately predict the impact of a mutagenesis experiment on the biosensor characteristics (dynamic range, operational range. . .). Structural approaches are probably the key to this issue and the recent development of structural prediction and docking programs may soon provide researchers with this type of tool (Mullard, 2021). Until recently, structure-based design was simply impossible in the absence of any relevant homology. With the advent of AI-based approaches to protein-structure prediction [e.g., AlphaFold (Jumper et al., 2021)]—The success stories of which suggest that *in silico* models at almost atomic resolution might be within reach—It might be possible to take the computational design of bespoke binding sites from merely instructing mutagenesis experiments on an, at best, semi-quantitative basis, to predictive reliability.

As shown for several applications in this review, biosensor circuits can be used to screen for the best performers in a bioproduction experiment by linking the reporter expression to the production level. High-throughput FACS or the less-scalable but more affordable plate-reader-based approach can be used to select the cells with higher reporter expression levels. However, this can become a problem if so-called cheaters come into play (Trivedi et al., 2022), giving high reporter signals without showing high production rates of the target compound. More methods that address this issue are needed. Finally, the most important future milestone for the field would be to develop a toolbox that can predict and model—Within the accuracy levels required for the application—The behaviour of a biosensor for a given design and target molecule, while taking process conditions and possible environmental perturbations into account that may appear during the experiment.

Author contributions

JT-L collected the biosensors list. MS wrote the structural biology part. All the authors wrote and approved the manuscript.

Funding

JT-L was supported by European Union Marie Skłodowska-Curie Action Individual Postdoctoral Fellowship (Grant agreement ID: 101062593). JT-L was supported by the Next-Generation EU (NGEU) fund through the Spanish Recovery, Transformation and Resilience Plan *via* a Margarita Salas personal grant from the Spanish Ministry of Universities (UNI/551/2021). MS was supported by the Next-Generation EU (NGEU) fund through the Spanish Recovery, Transformation and Resilience Plan *via* a María Zambrano personal grant from the Spanish Ministry of Universities (UNI/551/2021). PC acknowledges MCIN/AEI/10.13039/501100011033 funding through PID 2020-117271RB-C2 (BIODYNAMICS). PC was supported by the Spanish Ministry of Universities (UNI/551/2021), grant number UP 2021-036 funded by European Union - Next-generation EU. This research received financial support from Generalitat Valenciana through grant CIAICO/2021/159 (SmartBioFab). PC acknowledges MCIN/AEI/10.13039/501100011033 and NextGenerationEU/PRTR funding through grant TED 2021-131049 B-I00 (BioEcoDBTL). PC acknowledges MCIN/AEI/10.13039/501100011033 funding through grant PID 2021-127888NA-I00 (COMPSYNBIO).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alves, J. A., Previato-Mello, M., Barroso, K. C. M., Koide, T., and da Silva Neto, J. F. (2021). The MarR family regulator OsbR controls oxidative stress response, anaerobic nitrate respiration, and biofilm formation in Chromobacterium violaceum. *BMC Microbiol.* 21, 304. doi:10.1186/s12866-021-02369-x

Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjørn, S. P., Givskov, M., and Molin, S. (1998). New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol.* 64, 2240–2246. doi:10.1128/AEM.64.6.2240-2246.1998

Atkinson, J. T., Su, L., Zhang, X., Bennett, G. N., Silberg, J. J., and Ajo-Franklin, C. M. (2022). Real-time bioelectronic sensing of environmental contaminants. *Nature* 611, 548–553. doi:10.1038/s41586-022-05356-y

Baek, M., McHugh, R., Anishchenko, I., Baker, D., and DiMaio, F. (2022). Accurate prediction of nucleic acid and protein-nucleic acid complexes using RoseTTAFoldNA. *Bioinformatics*. doi:10.1101/2022.09.09.507333

Berepiki, A., Kent, R., Machado, L. F. M., and Dixon, N. (2020). Development of high-performance whole cell biosensors aided by statistical modeling. *ACS Synth. Biol.* 9, 576–589. doi:10.1021/acssynbio.9b00448

Boada, Y., Vignoni, A., Picó, J., and Carbonell, P. (2020). Extended metabolic biosensor design for dynamic pathway regulation of cell factories. *IScience* 23, 101305. doi:10.1016/j.isci.2020.101305

Bordoy, A. E., O'Connor, N. J., and Chatterjee, A. (2019). Construction of two-input logic gates using transcriptional interference. *ACS Synth. Biol.* 8, 2428–2441. doi:10.1021/acssynbio.9b00321

Brödel, A. K., Rodrigues, R., Jaramillo, A., and Isalan, M. (2020). Accelerated evolution of a minimal 63–amino acid dual transcription factor. *Sci. Adv.* 6, eaba2728. doi:10.1126/sciadv.aba2728

Cadet, F., Fontaine, N., Li, G., Sanchis, J., Ng, F., Chong, M., et al. (2018). A machine learning approach for reliable prediction of amino acid interactions and its application in the directed evolution of enantioselective enzymes. *Sci. Rep.* 8, 16757. doi:10.1038/s41598-018-35033-y

Cann, O. (2016). *These are the top 10 emerging technologies of 2016*. Cologny, Switzerland: Word Economic Forum. Available at: https://www3.weforum.org/docs/GAC16_Top10_Emerging_Technologies_2016_report.pdf

Carbajosa, G., Trigo, A., Valencia, A., and Cases, I. (2009). Bionemo: Molecular information on biodegradation metabolism. *Nucleic Acids Res.* 37, D598–D602. doi:10.1093/nar/gkn864

Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., et al. (2018). An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun. Biol.* 1, 66. doi:10.1038/s42003-018-0076-9

Carbonell, P., Wong, J., Swainston, N., Takano, E., Turner, N. J., Scrutton, N. S., et al. (2018). Selenzyme: Enzyme selection tool for pathway design. *Bioinformatics* 34, 2153–2154. doi:10.1093/bioinformatics/bty065

Carpenter, A., Paulsen, I., and Williams, T. (2018). Blueprints for biosensors: Design, limitations, and applications. *Genes* 9, 375. doi:10.3390/genes9080375

Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., et al. (2022). Jaspar 2022: The 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 50, D165–D173. doi:10.1093/nar/gkab1113

Chen, C., Hou, J., Shi, X., Yang, H., Birchler, J. A., and Cheng, J. (2021). DeepGRN: Prediction of transcription factor binding site across cell-types using attention-based deep neural networks. *BMC Bioinforma.* 22, 38. doi:10.1186/s12859-020-03952-1

Chen, Y., Ho, J. M. L., Shis, D. L., Gupta, C., Long, J., Wagner, D. S., et al. (2018). Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nat. Commun.* 9, 64. doi:10.1038/s41467-017-02473-5

Choi, O., Lee, Y., Han, I., Kim, H., Goo, E., Kim, J., et al. (2013). A simple and sensitive biosensor strain for detecting toxoflavin using β-galactosidase activity. *Biosens. Bioelectron.* 50, 256–261. doi:10.1016/j.bios.2013.06.058

Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006). Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nat. Biotechnol.* 24, 708–712. doi:10.1038/nbt1209

Corso, G., Stärk, H., Jing, B., Barzilay, R., and Jaakkola, T. (2022). DiffDock: Diffusion steps, twists, and turns for molecular docking. Available at: https://arxiv.org/abs/2210.01776.doi:10.48550/ARXIV.2210.01776

De Paepe, B., Maertens, J., Vanholme, B., and De Mey, M. (2019). Chimeric LysR-type transcriptional biosensors for customizing ligand specificity profiles toward flavonoids. *ACS Synth. Biol.* 8, 318–331. doi:10.1021/acssynbio.8b00326

De Paepe, B., Maertens, J., Vanholme, B., and De Mey, M. (2018). Modularization and response curve engineering of a naringenin-responsive transcriptional biosensor. *ACS Synth. Biol.* 7, 1303–1314. doi:10.1021/acssynbio.7b00419

De Paepe, B., Peters, G., Coussement, P., Maertens, J., and De Mey, M. (2017). Tailor-made transcriptional biosensors for optimizing microbial cell factories. *J. Ind. Microbiol. Biotechnol.* 44, 623–645. doi:10.1007/s10295-016-1862-3

Delépine, B., Libis, V., Carbonell, P., and Faulon, J-L. (2016). SensiPath: Computer-aided design of sensing-enabling metabolic pathways. *Nucleic Acids Res.* 44, W226–W231. doi:10.1093/nar/gkw305

Della Corte, D., van Beek, H. L., Syberg, F., Schallmey, M., Tobola, F., Cormann, K. U., et al. (2020). Engineering and application of a biosensor with focused ligand specificity. *Nat. Commun.* 11, 4851. doi:10.1038/s41467-020-18400-0

Deplancke, B., and Gheldof, N. (2012). *Gene regulatory networks: Methods and protocols*. New York: Humana Press ; Springer.

d'Oelsnitz, S., Kim, W., Burkholder, N. T., Javanmardi, K., Thyer, R., Zhang, Y., et al. (2022). Using fungible biosensors to evolve improved alkaloid biosyntheses. *Nat. Chem. Biol.* 18, 981–989. doi:10.1038/s41589-022-01072-w

Doong, S. J., Gupta, A., and Prather, K. L. J. (2018). Layered dynamic regulation for improving metabolic pathway productivity in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 115, 2964–2969. doi:10.1073/pnas.1716920115

Dudek, C-A., and Jahn, D. (2022). Prodoric: State-of-the-art database of prokaryotic gene regulation. *Nucleic Acids Res.* 50, D295–D302. doi:10.1093/ar/gkab1110

Eichner, J., Topf, F., Dräger, A., Wrzodek, C., Wanke, D., and Zell, A. (2013). TFpredict and SABINE: Sequence-based prediction of structural and functional characteristics of transcription factors. *PLoS ONE* 8, e82238. doi:10.1371/journal.pone.0082238

Ellefson, J. W., Ledbetter, M. P., and Ellington, A. D. (2018). Directed evolution of a synthetic phylogeny of programmable Trp repressors. *Nat. Chem. Biol.* 14, 361–367. doi:10.1038/s41589-018-0006-7

Ellefson, J. W., Meyer, A. J., Hughes, R. A., Cannon, J. R., Brodbelt, J. S., and Ellington, A. D. (2014). Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol.* 32, 97–101. doi:10.1038/nbt.2714

Esvelt, K. M., Carlson, J. C., and Liu, D. R. (2011). A system for the continuous directed evolution of biomolecules. *Nature* 472, 499–503. doi:10.1038/nature09929

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., et al. (2021). Protein complex prediction with AlphaFold-Multimer. *Bioinformatics*. doi:10.1101/2021.10.04.463034

Fernandez-López, R., Ruiz, R., de la Cruz, F., and Moncalián, G. (2015). Transcription factor-based biosensors enlightened by the analyte. *Front. Microbiol.* 6, 648. doi:10.3389/fmicb.2015.00648

Gao, J., Du, M., Zhao, J., zhang, Y., Xu, N., Du, H., et al. (2022). Design of a genetically encoded biosensor to establish a high-throughput screening platform for L-cysteine overproduction. *Metab. Eng.* 73, 144–157. doi:10.1016/j.ymben.2022.07.007

Gearing, L. J., Cumming, H. E., Chapman, R., Finkel, A. M., Woodhouse, I. B., Luu, K., et al. (2019). CiiiDER: A tool for predicting and analysing transcription factor binding sites. *PLOS ONE* 14, e0215495. doi:10.1371/journal.pone.0215495

Golitsch, F., Bücking, C., and Gescher, J. (2013). Proof of principle for an engineered microbial biosensor based on Shewanella oneidensis outer membrane protein complexes. *Biosens. Bioelectron.* 47, 285–291. doi:10.1016/j.bios.2013.03.010

Grainger, D. C., Aiba, H., Hurd, D., Browning, D. F., and Busby, S. J. W. (2007). Transcription factor distribution in *Escherichia coli*: Studies with FNR protein. *Nucleic Acids Res.* 35, 269–278. doi:10.1093/nar/gkl1023

Greener, J. G., and Sternberg, M. J. (2015). AlloPred: Prediction of allosteric pockets on proteins using normal mode perturbation analysis. *BMC Bioinforma.* 16, 335. doi:10.1186/s12859-015-0771-1

Guo, Y., Hui, C., Liu, L., Chen, M., and Huang, H. (2021). Development of a bioavailable Hg(II) sensing system based on MerR-regulated visual pigment biosynthesis. *Sci. Rep.* 11, 13516. doi:10.1038/s41598-021-92878-6

Gurevich, I., Zhang, C., and Aneskievich, B. J. (2010). Scanning for transcription factor binding by a variant EMSA. *Methods Mol. Biol.* 585, 147–158. doi:10.1007/978-1-60761-380-0_11

Hanko, E. K. R., Paiva, A. C., Jonczyk, M., Abbott, M., Minton, N. P., and Malys, N. (2020). A genome-wide approach for identification and characterisation of metabolite-inducible systems. *Nat. Commun.* 11, 1213. doi:10.1038/s41467-020-14941-6

Hansen, M. L., He, Z., Wibowo, M., and Jelsbak, L. (2021). A whole-cell biosensor for detection of 2,4-diacetylphloroglucinol (DAPG)-Producing bacteria from grassland soil. *Appl. Environ. Microbiol.* 87, e01400–e01420. doi:10.1128/AEM.01400-20

Hartline, C. J., Schmitz, A. C., Han, Y., and Zhang, F. (2021). Dynamic control in metabolic engineering: Theories, tools, and applications. *Metab. Eng.* 63, 126–140. doi:10.1016/j.ymben.2020.08.015

Hilpert, C., Beranger, L., Souza, P. C. T., Vainikka, P. A., Nieto, V., Marrink, S. J., et al. (2022). Facilitating CG simulations with MAD: The MArtini database server. *Biophysics*. doi:10.1101/2022.08.03.502585

Hodgman, C. E., and Jewett, M. C. (2012). Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng.* 14, 261–269. doi:10.1016/j.ymben.2011.09.002

Hossain, A., Lopez, E., Halper, S. M., Cetnar, D. P., Reis, A. C., Strickland, D., et al. (2020). Automated design of thousands of nonrepetitive parts for engineering stable genetic systems. *Nat. Biotechnol.* 38, 1466–1475. doi:10.1038/s41587-020-0584-2

Hu, H., Miao, Y-R., Jia, L-H., Yu, Q-Y., Zhang, Q., and Guo, A-Y. (2019). AnimalTFDB 3.0: A comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res.* 47, D33–D38. doi:10.1093/nar/gky822

Huang, C-H., Shen, C. R., Li, H., Sung, L-Y., Wu, M-Y., and Hu, Y-C. (2016). CRISPR interference (CRISPRi) for gene regulation and succinate production in cyanobacterium S. elongatus PCC 7942. *Microb. Cell Factories* 15, 196. doi:10.1186/s12934-016-0595-3

Huang, M., Song, K., Liu, X., Lu, S., Shen, Q., Wang, R., et al. (2018). AlloFinder: A strategy for allosteric modulator discovery and allosterome analyses. *Nucleic Acids Res.* 46, W451–W458. doi:10.1093/nar/gky374

Hui, C., Guo, Y., Liu, L., Zhang, N., Gao, C., Yang, X., et al. (2020). Genetic control of violacein biosynthesis to enable a pigment-based whole-cell lead biosensor. *RSC Adv.* 10, 28106–28113. doi:10.1039/D0RA04815A

Jin, J., Tian, F., Yang, D-C., Meng, Y-Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi:10.1093/nar/gkw982

Juárez, J. F., Lecube-Azpeitia, B., Brown, S. L., Johnston, C. D., and Church, G. M. (2018). Biosensor libraries harness large classes of binding domains for construction of allosteric transcriptional regulators. *Nat. Commun.* 9, 3101. doi:10.1038/s41467-018-05525-6

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2

Jung, J. K., Alam, K. K., Verosloff, M. S., Capdevila, D. A., Desmau, M., Clauer, P. R., et al. (2020). Cell-free biosensors for rapid detection of water contaminants. *Nat. Biotechnol.* 38, 1451–1459. doi:10.1038/s41587-020-0571-7

Kaczmarek, J. A., and Prather, K. L. J. (2021). Effective use of biosensors for high-throughput library screening for metabolite production. *J. Ind. Microbiol. Biotechnol.* 48, kuab049. doi:10.1093/jimb/kuab049

Kılıç, S., White, E. R., Sagitova, D. M., Cornish, J. P., and Erill, I. (2014). CollecTF: A database of experimentally validated transcription factor-binding sites in bacteria. *Nucleic Acids Res.* 42, D156–D160. doi:10.1093/nar/gkt1123

Kim, G. B., Gao, Y., Palsson, B. O., and Lee, S. Y. (2021). DeepTFactor: A deep learning-based tool for the prediction of transcription factors. *Proc. Natl. Acad. Sci.* 118, e2021171118. doi:10.1073/pnas.2021171118

Kim, J. Y. H., and Cha, H. J. (2003). Down-regulation of acetate pathway through antisense strategy inEscherichia coli: Improved foreign protein production. *Biotechnol. Bioeng.* 83, 841–853. doi:10.1002/bit.10735

Koch, M., Pandi, A., Delépine, B., and Faulon, J-L. (2018). A dataset of small molecules triggering transcriptional and translational cellular responses. *Data Brief.* 17, 1374–1378. doi:10.1016/j.dib.2018.02.061

Koo, P. K., and Ploenzke, M. (2020). Deep learning for inferring transcription factor binding sites. *Curr. Opin. Syst. Biol.* 19, 16–23. doi:10.1016/j.coisb.2020.04.001

Li, G., Du, X., Li, X., Zou, L., Zhang, G., and Wu, Z. (2021). Prediction of DNA binding proteins using local features and long-term dependencies with primary sequences based on deep learning. *PeerJ* 9, e11262. doi:10.7717/peerj.11262

Li, H., Chen, W., Jin, R., Jin, J-M., and Tang, S-Y. (2017). Biosensor-aided high-throughput screening of hyper-producing cells for malonyl-CoA-derived products. *Microb. Cell Factories* 16, 187. doi:10.1186/s12934-017-0794-6

Lin, G-M., Warden-Rothman, R., and Voigt, C. A. (2019). Retrosynthetic design of metabolic pathways to chemicals not found in nature. *Curr. Opin. Syst. Biol.* 14, 82–107. doi:10.1016/j.coisb.2019.04.004

Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., et al. (2022). Evolutionary-scale prediction of atomic level protein structure with a language model. *Synth. Biol.* doi:10.1101/2022.07.20.500902

Liu, D., and Zhang, F. (2018). Metabolic feedback circuits provide rapid control of metabolite dynamics. *ACS Synth. Biol.* 7, 347–356. doi:10.1021/acssynbio.7b00342

Liu, Y., Li, Q., Zheng, P., Zhang, Z., Liu, Y., Sun, C., et al. (2015). Developing a high-throughput screening method for threonine overproduction based on an artificial promoter. *Microb. Cell Factories* 14, 121. doi:10.1186/s12934-015-0311-8

Liu, Y., Liu, Y., and Wang, M. (2017). Design, optimization and application of small molecule biosensor in metabolic engineering. *Front. Microbiol.* 8, 2012. doi:10.3389/fmicb.2017.02012

Lopreside, A., Wan, X., Michelini, E., Roda, A., and Wang, B. (2019). Comprehensive profiling of diverse genetic reporters with application to whole-cell and cell-free biosensors. *Anal. Chem.* 91, 15284–15292. doi:10.1021/acs.analchem.9b04444

Lucas, J. E., and Kortemme, T. (2020). New computational protein design methods for de novo small molecule binding sites. *PLOS Comput. Biol.* 16, e1008178. doi:10.1371/journal.pcbi.1008178

Machado, F. M. L., Currin, A., and Dixon, N. (2019). Directed evolution of the PcaV allosteric transcription factor to generate a biosensor for aromatic aldehydes. *J. Biol. Eng.* 13, 91. doi:10.1186/s13036-019-0214-z

Machado, L. F. M., and Dixon, N. (2022). Directed evolution of transcription factor-based biosensors for altered effector specificity. *Methods Mol. Biol.* 2461, 175–193. doi:10.1007/978-1-0716-2152-3_12

Mannan, A. A., Liu, D., Zhang, F., and Oyarzún, D. A. (2017). Fundamental design principles for transcription-factor-based metabolite biosensors. *ACS Synth. Biol.* 6, 1851–1859. doi:10.1021/acssynbio.7b00172

Meng, X., Brodsky, M. H., and Wolfe, S. A. (2005). A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nat. Biotechnol.* 23, 988–994. doi:10.1038/nbt1120

Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J., and Alba, M. M. (2002). Promo: Detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* 18, 333–334. doi:10.1093/bioinformatics/18.2.333

Miller, C. A., Ho, J. M. L., and Bennett, M. R. (2022). Strategies for improving small-molecule biosensors in bacteria. *Biosensors* 12, 64. doi:10.3390/bios12020064

Mishra, A., Pokhrel, P., and Hoque, M. T. (2019). StackDPPred: A stacking based prediction of DNA-binding protein from sequence. *Bioinformatics* 35, 433–441. doi:10.1093/bioinformatics/bty653

Modrzejewska, M., Kawalek, A., and Bartosik, A. A. (2021). The lrp/AsnC-type regulator PA2577 controls the EamA-like transporter gene PA2576 in *Pseudomonas aeruginosa*. *Int. J. Mol. Sci.* 22, 13340. doi:10.3390/ijms222413340

Moraskie, M., Roshid, M. H. O., O'Connor, G., Dikici, E., Zingg, J-M., Deo, S., et al. (2021). Microbial whole-cell biosensors: Current applications, challenges, and future perspectives. *Biosens. Bioelectron.* 191, 113359. doi:10.1016/j.bios.2021.113359

Mullard, A. (2021). What does AlphaFold mean for drug discovery? *Nat. Rev. Drug Discov.* 20, 725–727. doi:10.1038/d41573-021-00161-0

Naidu, R., Biswas, B., Willett, I. R., Cribb, J., Kumar Singh, B., Paul Nathanail, C., et al. (2021). Chemical pollution: A growing peril and potential catastrophic risk to humanity. *Environ. Int.* 156, 106616. doi:10.1016/j.envint.2021.106616

Neubauer, P., and Junne, S. (2010). Scale-down simulators for metabolic analysis of large-scale bioprocesses. *Curr. Opin. Biotechnol.* 21, 114–121. doi:10.1016/j.copbio.2010.02.001

Nielsen, A. A. K., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., et al. (2016). Genetic circuit design automation. *Science* 352, aac7341. doi:10.1126/science.aac7341

Nijman, S., and Halpaap, A. (2019). *UN report: Urgent action needed to tackle chemical pollution as global production is set to double by 2030*. Nairobi, Kenya: United Nations Environment Programme.

Nourmohammadi, E., Hosseinkhani, S., Nedaeinia, R., Khoshdel-Sarkarizi, H., Nedaeinia, M., Ranjbar, M., et al. (2020). Construction of a sensitive and specific lead biosensor using a genetically engineered bacterial system with a luciferase gene reporter controlled by pbr and cadA promoters. *Biomed. Eng. OnLine* 19, 79. doi:10.1186/s12938-020-00816-w

Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., et al. (2013). RegPrecise 3.0 – a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* 14, 745. doi:10.1186/1471-2164-14-745

Oliveira Monteiro, L. M., Saraiva, J., Toscan, R. B., Stadler, P. F., Silva-Rocha, R., and da Rocha, U. N. (2021). PredicTF: A tool to predict bacterial transcription factors in complex microbial communities. *Bioinformatics*. doi:10.1101/2021.01.28.428666

Orsi, E., Claassens, N. J., Nikel, P. I., and Lindner, S. N. (2021). Growth-coupled selection of synthetic modules to accelerate cell factory development. *Nat. Commun.* 12, 5295. doi:10.1038/s41467-021-25665-6

Ortet, P., De Luca, G., Whitworth, D. E., and Barakat, M. (2012). P2TF: A comprehensive resource for analysis of prokaryotic transcription factors. *BMC Genomics* 13, 628. doi:10.1186/1471-2164-13-628

Pandi, A., Grigoras, I., Borkowski, O., and Faulon, J-L. (2019). Optimizing cell-free biosensors to monitor enzymatic production. *ACS Synth. Biol.* 8, 1952–1957. doi:10.1021/acssynbio.9b00160

Pandi, A., Koch, M., Voyvodic, P. L., Soudier, P., Bonnet, J., Kushwaha, M., et al. (2019). Metabolic perceptrons for neural computing in biological systems. *Nat. Commun.* 10, 3880. doi:10.1038/s41467-019-11889-0

Panjkovich, A., and Daura, X. (2014). Pars: A web server for the prediction of protein allosteric and regulatory sites. *Bioinformatics* 30, 1314–1315. doi:10.1093/bioinformatics/btu002

Pearce, R., Omenn, G. S., and ZhangNovo, Y. D. (2022). RNA tertiary structure prediction at atomic resolution using geometric potentials from deep learning. *Bioinformatics*. doi:10.1101/2022.05.15.491755

Polizzi, N. F., and DeGrado, W. F. (2020). A defined structural unit enables de novo design of small-molecule–binding proteins. *Science* 369, 1227–1233. doi:10.1126/science.abb8330

Pujato, M., Kieken, F., Skiles, A. A., Tapinos, N., and Fiser, A. (2014). Prediction of DNA binding motifs from 3D models of transcription factors; identifying TLX3 regulated genes. *Nucleic Acids Res.* 42, 13500–13512. doi:10.1093/nar/gku1228

Rajput, A., Kaur, K., and Kumar, M. (2016). SigMol: Repertoire of quorum sensing signaling molecules in prokaryotes. *Nucleic Acids Res.* 44, D634–D639. doi:10.1093/nar/gkv1076

Rondon, R. E., Groseclose, T. M., Short, A. E., and Wilson, C. J. (2019). Transcriptional programming using engineered systems of transcription factors and genetic architectures. *Nat. Commun.* 10, 4784. doi:10.1038/s41467-019-12706-4

Rottinghaus, A. G., Xi, C., Amrofell, M. B., Yi, H., and Moon, T. S. (2022). Engineering ligand-specific biosensors for aromatic amino acids and neurochemicals. *Cell Syst.* 13, 204–214.e4. doi:10.1016/j.cels.2021.10.006

Saito, Y., Oikawa, M., Nakazawa, H., Niide, T., Kameda, T., Tsuda, K., et al. (2018). Machine-learning-Guided mutagenesis for directed evolution of fluorescent proteins. *ACS Synth. Biol.* 7, 2014–2022. doi:10.1021/acssynbio.8b00155

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sánchez-Pérez, M., Gómez-Romero, L., Ledezma-Tejeida, D., et al. (2019). RegulonDB v 10.5: Tackling challenges to unify classic and high throughput knowledge of gene regulation in*E. coli*K-12 in *E. coli* K-12. *Nucleic Acids Res.* 47, D212–D220. D212–20. doi:10.1093/nar/gky1077

Sebastian, A., and Contreras-Moreira, B. (2014). footprintDB: a database of transcription factors with annotated cis elements and binding interfaces. *Bioinformatics* 30, 258–265. doi:10.1093/bioinformatics/btt663

Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., et al. (2021). Mol* viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.* 49, W431–W437. doi:10.1093/nar/gkab314

Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat. Methods* 2, 905–909. doi:10.1038/nmeth819

Siedler, S., Khatri, N. K., Zsohár, A., Kjærbølling, I., Vogt, M., Hammar, P., et al. (2017). Development of a bacterial biosensor for rapid screening of yeast *p*-coumaric acid production P-coumaric acid production. *ACS Synth. Biol.* 6, 1860–1869. doi:10.1021/acssynbio.7b00009

Snoek, T., Chaberski, E. K., Ambri, F., Kol, S., Bjørn, S. P., Pang, B., et al. (2020). Evolution-guided engineering of small-molecule biosensors. *Nucleic Acids Res.* 48, e3. doi:10.1093/nar/gkz954

Sonntag, C. K., Flachbart, L. K., Maass, C., Vogt, M., and Marienhagen, J. (2020). A unified design allows fine-tuning of biosensor parameters and application across bacterial species. *Metab. Eng. Commun.* 11, e00150. doi:10.1016/j.mec.2020.e00150

Soudier, P., Faure, L., Kushwaha, M., and Faulon, J-L. (2022). Cell-free biosensors and AI integration. *Methods Mol. Biol.* 2433, 303–323. doi:10.1007/978-1-0716-1998-8_19

Souza, P. C. T., Thallmair, S., Conflitti, P., Ramírez-Palacios, C., Alessandri, R., Raniolo, S., et al. (2020). Protein–ligand binding with the coarse-grained Martini model. *Nat. Commun.* 11, 3714. doi:10.1038/s41467-020-17437-5

Stevens, J. T., and Carothers, J. M. (2015). Designing RNA-based genetic control systems for efficient production from engineered metabolic pathways. *ACS Synth. Biol.* 4, 107–115. doi:10.1021/sb400201u

Su, Y., and Hammond, M. C. (2020). RNA-based fluorescent biosensors for live cell imaging of small molecules and RNAs. *Curr. Opin. Biotechnol.* 63, 157–166. doi:10.1016/j.copbio.2020.01.001

Tellechea-Luzardo, J., Otero-Muras, I., Goñi-Moreno, A., and Carbonell, P. (2022). Fast biofoundries: Coping with the challenges of biomanufacturing. *Trends Biotechnol.* 40, 831–842. doi:10.1016/j.tibtech.2021.12.006

Teng, Y., Zhang, J., Jiang, T., Zou, Y., Gong, X., and Yan, Y. (2022). Biosensor-enabled pathway optimization in metabolic engineering. *Curr. Opin. Biotechnol.* 75, 102696. doi:10.1016/j.copbio.2022.102696

Tenhaef, N., Stella, R., Frunzke, J., and Noack, S. (2021). Automated rational strain construction based on high-throughput conjugation. *ACS Synth. Biol.* 10, 589–599. doi:10.1021/acssynbio.0c00599

Trivedi, V. D., Mohan, K., Chappell, T. C., Mays, Z. J. S., and Nair, N. U. (2022). Cheating the cheater: Suppressing false-positive enrichment during biosensor-guided biocatalyst engineering. *ACS Synth. Biol.* 11, 420–429. doi:10.1021/acssynbio.1c00506

Umeno, D., Kimura, Y., and Kawai-Noma, S. (2021). Transcription factors as evolvable biosensors. *Anal. Sci.* 37, 699–703. doi:10.2116/analsci.20SCR12

van Werven, F. J. (2006). The use of biotin tagging in *Saccharomyces cerevisiae* improves the sensitivity of chromatin immunoprecipitation. *Nucleic Acids Res.* 34, e33. doi:10.1093/nar/gkl003

Volk, M. J., Lourentzou, I., Mishra, S., Vo, L. T., Zhai, C., and Zhao, H. (2020). Biosystems design by machine learning. *ACS Synth. Biol.* 9, 1514–1533. doi:10.1021/acssynbio.0c00129

Wang, J., Li, C., Jiang, T., and Yan, Y. (2023). Biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening for over-production phenotypes. *Metab. Eng.* 75, 58–67. doi:10.1016/j.ymben.2022.11.004

Wang, R., Cress, B. F., Yang, Z., Hordines, J. C., Zhao, S., Jung, G. Y., et al. (2019). Design and characterization of biosensors for the screening of modular assembled naringenin biosynthetic library in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 8, 2121–2130. doi:10.1021/acssynbio.9b00212

Wang, T., Sun, W., Fan, L., Hua, C., Wu, N., Fan, S., et al. (2021). An atlas of the binding specificities of transcription factors in *Pseudomonas aeruginosa* directs prediction of novel regulators in virulence. *ELife* 10, e61885. doi:10.7554/eLife.61885

Wang, Z., Civelek, M., Miller, C. L., Sheffield, N. C., Guertin, M. J., and ZangBART, C. (2018). Bart: A transcription factor prediction tool with query gene sets or epigenomic profiles. *Bioinformatics* 34, 2867–2869. doi:10.1093/bioinformatics/bty194

Watstein, D. M., McNerney, M. P., and Styczynski, M. P. (2015). Precise metabolic engineering of carotenoid biosynthesis in *Escherichia coli* towards a low-cost biosensor. *Metab. Eng.* 31, 171–180. doi:10.1016/j.ymben.2015.06.007

Webster, D. P., TerAvest, M. A., Doud, D. F. R., Chakravorty, A., Holmes, E. C., Radens, C. M., et al. (2014). An arsenic-specific biosensor with genetically engineered Shewanella oneidensis in a bioelectrochemical system. *Biosens. Bioelectron.* 62, 320–324. doi:10.1016/j.bios.2014.07.003

Wehrs, M., Tanjore, D., Eng, T., Lievense, J., Pray, T. R., and Mukhopadhyay, A. (2019). Engineering robust production microbes for large-scale cultivation. *Trends Microbiol.* 27, 524–537. doi:10.1016/j.tim.2019.01.006

Wen, K. Y., Cameron, L., Chappell, J., Jensen, K., Bell, D. J., Kelwick, R., et al. (2017). A cell-free biosensor for detecting quorum sensing molecules in *P. aeruginosa*-infected respiratory samples in *P. aeruginosa* -infected respiratory samples. *ACS Synth. Biol.* 6, 2293–2301. doi:10.1021/acssynbio.7b00219

Wu, G., Yan, Q., Jones, J. A., Tang, Y. J., Fong, S. S., and Koffas, M. A. G. (2016). Metabolic burden: Cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* 34, 652–664. doi:10.1016/j.tibtech.2016.02.010

Wu, J., Jiang, P., Chen, W., Xiong, D., Huang, L., Jia, J., et al. (2017). Design and application of a lactulose biosensor. *Sci. Rep.* 7, 45994. doi:10.1038/srep45994

Wu, Y., Chen, T., Liu, Y., Tian, R., Lv, X., Li, J., et al. (2020). Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. *Nucleic Acids Res.* 48, 996–1009. doi:10.1093/nar/gkz1123

Wu, Z., Kan, S. B. J., Lewis, R. D., Wittmann, B. J., and Arnold, F. H. (2019). Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc. Natl. Acad. Sci.* 116, 8852–8858. doi:10.1073/pnas.1901979116

Xiao, S., Tian, H., and Tao, P. (2021). PASSer2.0: Accurate prediction of protein allosteric sites through automated machine learning. *Chemistry*. doi:10.26434/chemrxiv-2021-q4319-v2

Xu, P., Li, L., Zhang, F., Stephanopoulos, G., and Koffas, M. (2014). Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proc. Natl. Acad. Sci.* 111, 11299–11304. doi:10.1073/pnas.1406401111

Xu, X., Li, X., Liu, Y., Zhu, Y., Li, J., Du, G., et al. (2020). Pyruvate-responsive genetic circuits for dynamic control of central metabolism. *Nat. Chem. Biol.* 16, 1261–1268. doi:10.1038/s41589-020-0637-3

Xu, X., Ma, Z., Sun, H., and Zou, X. S. M-T. F. (2016). SM-TF: A structural database of small molecule-transcription factor complexes. *J. Comput. Chem.* 37, 1559–1564. doi:10.1002/jcc.24370

Yang, Y., Lin, Y., Wang, J., Wu, Y., Zhang, R., Cheng, M., et al. (2018). Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. *Nat. Commun.* 9, 3043. doi:10.1038/s41467-018-05466-0

Yilmaz, A., Nishiyama, M. Y., Fuentes, B. G., Souza, G. M., Janies, D., Gray, J., et al. (2009). Grassius: A platform for comparative regulatory genomics across the grasses. *Plant Physiol.* 149, 171–180. doi:10.1104/pp.108.128579

Yoshida, K., Inoue, K., Takahashi, Y., Ueda, S., Isoda, K., Yagi, K., et al. (2008). Novel carotenoid-based biosensor for simple visual detection of arsenite: Characterization and preliminary evaluation for environmental application. *Appl. Environ. Microbiol.* 74, 6730–6738. doi:10.1128/AEM.00498-08

Zeng, Z., Shi, H., Wu, Y., and Hong, Z. (2015). Survey of Natural Language Processing techniques in bioinformatics. *Comput. Math. Methods Med.* 2015, 1–10. doi:10.1155/2015/674296

Zhang, Y-H. P., Sun, J., and Ma, Y. (2017). Biomanufacturing: History and perspective. *J. Ind. Microbiol. Biotechnol.* 44, 773–784. doi:10.1007/s10295-016-1863-2

Zheng, S., Hou, J., Zhou, Y., Fang, H., Wang, T-T., Liu, F., et al. (2018). One-pot two-strain system based on glucaric acid biosensor for rapid screening of myo-inositol oxygenase mutations and glucaric acid production in recombinant cells. *Metab. Eng.* 49, 212–219. doi:10.1016/j.ymben.2018.08.005

Zhou, A. Y., Baruch, M., Ajo-Franklin, C. M., and Maharbiz, M. M. (2017). A portable bioelectronic sensing system (BESSY) for environmental deployment incorporating differential microbial sensing in miniaturized reactors. *PLOS ONE* 12, e0184994. doi:10.1371/journal.pone.0184994

Zhou, S., Yuan, S-F., Nair, P. H., Alper, H. S., Deng, Y., and Zhou, J. (2021). Development of a growth coupled and multi-layered dynamic regulation network balancing malonyl-CoA node to enhance (2S)-naringenin biosynthesis in *Escherichia coli*. *Metab. Eng.* 67, 41–52. doi:10.1016/j.ymben.2021.05.007

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 07 February 2023
DOI 10.3389/fbioe.2023.1130368

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Jian-Wen Ye,
South China University of Technology, China
Quanfeng Liang,
Shandong University, China

*CORRESPONDENCE
Li-Hai Fan,
fanlh@fzu.edu.cn
Zheng-Jun Li,
lizj@mail.buct.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 23 December 2022
ACCEPTED 30 January 2023
PUBLISHED 07 February 2023

CITATION
Li H-F, Tian L, Lian G, Fan L-H and Li Z-J (2023), Engineering *Vibrio alginolyticus* as a novel chassis for PHB production from starch.
Front. Bioeng. Biotechnol. 11:1130368.
doi: 10.3389/fbioe.2023.1130368

Engineering *Vibrio alginolyticus* as a novel chassis for PHB production from starch

Hong-Fei Li[1,2,3], Linyue Tian[2], Guoli Lian[2], Li-Hai Fan[1,3]* and Zheng-Jun Li[2]*

[1]College of Chemical Engineering, Fujian Engineering Research Center of Advanced Manufacturing Technology for Fine Chemicals, Fuzhou University, Fuzhou, China, [2]College of Life Science and Technology, Beijing University of Chemical Technology, Beijing, China, [3]Qingyuan Innovation Laboratory, Quanzhou, China

Vibrio alginolyticus LHF01 was engineered to efficiently produce poly-3-hydroxybutyrate (PHB) from starch in this study. Firstly, the ability of *Vibrio alginolyticus* LHF01 to directly accumulate PHB using soluble starch as the carbon source was explored, and the highest PHB titer of 2.06 g/L was obtained in 18 h shake flask cultivation. Then, with the analysis of genomic information of *V. alginolyticus* LHF01, the PHB synthesis operon and amylase genes were identified. Subsequently, the effects of overexpressing PHB synthesis operon and amylase on PHB production were studied. Especially, with the co-expression of PHB synthesis operon and amylase, the starch consumption rate was improved and the PHB titer was more than doubled. The addition of 20 g/L insoluble corn starch could be exhausted in 6-7 h cultivation, and the PHB titer was 4.32 g/L. To the best of our knowledge, *V. alginolyticus* was firstly engineered to produce PHB with the direct utilization of starch, and this stain can be considered as a novel host to produce PHB using starch as the raw material.

KEYWORDS

amylase, metabolic engineering, poly-3-hydroxybutyrate, starch, *Vibrio alginolyticus*

1 Introduction

With the development of modern biotechnology and synthetic biology, microorganisms have been considered as important chassis cells for the production of biological products. While rational metabolic engineering of microorganisms greatly reduces the cost of production, the time required for cell growth and product synthesis during fermentation process is likely to be a potential limitation (Hoff et al., 2020). As is all known, *Escherichia coli* has been extensively studied as the host bacterium, but the range and efficiency of the biosynthesis applications begin to be limited due to its intrinsic capabilities. Interestingly, *Vibrio* strains are gaining more and more attentions due to their rapid growth rates. For example, the doubling time of *Vibrio alginolyticus* and *Vibrio parahaemolyticus* is reported to be 12-14 min (Hoff et al., 2020; Li et al., 2021). Another bacterium *Vibrio natriegens* is reported to possess the fastest cell growth rate of known bacteria, with a doubling time of less than 10 min (Weinstock et al., 2016). Besides, the genomic editing tools of *Vibrio* have also been gradually established (Roh et al., 2012; Dalia et al., 2017; Lee et al., 2019). Therefore, the metabolic engineering of *Vibrio* strains as host bacteria for biological production would provide a new direction for the development of synthetic biology.

Polyhydroxyalkanoate (PHA) is an environment-friendly alternative to replace petroleum-based plastics due to its biodegradability, and the most common of which is

poly-3-hydroxybutyrate (PHB) (Zheng et al., 2020). Generally, PHA is accumulated as intracellular lipid inclusion bodies to store carbon and energy in microbial cells (Choi et al., 2020; Pagliano et al., 2021). Up to now, there are relatively few studies on the PHA production using *Vibrio* strains. A few research indicated that *Vibrio* strain are responsible for PHA production in mixed microbial cultures (Cui et al., 2016; Zhao et al., 2021). Recently, *V. alginolyticus* and *V. proteolyticus* were reported as novel candidates for PHA production (Hong et al., 2019; Li et al., 2021). Studies have shown that the biosynthesis of PHA is a complex metabolic process, which involved a variety of enzymes (Meng et al., 2014). The most common pathway can be called the "three-step synthesis" pathway, in which the carbon source is converted into acetyl-CoA by microbial decomposition, and then two acetyl-CoA molecules are condensed into acetoacetyl-CoA by acetyl-CoA acetyltransferase (PhaA). Subsequently, 3-hydroxybutyryl-CoA is generated through the reduction of acetoacetyl-CoA under the catalysis of NADPH-dependent reductase (PhaB). Finally, PHA synthase (PhaC) polymerize the 3-hydroxybutyryl-CoA monomer to generate PHB (Sagong et al., 2018). There are also some other enzymes that can regulate PHA biosynthesis, such as the phasin family protein (PhaP) (Lee et al., 2021).

High production cost is a great limitation for the application of PHA as biodegradable materials (Liu et al., 2021). Industrial and agricultural wastes, such as volatile fatty acids (Pu et al., 2020a; Wang et al., 2022), waste potato starch (Haas et al., 2008), wheat straw (Dahman and Ugwu, 2014), and waste frying oil (Tian et al., 2022) have been proposed as inexpensive substrates to reduce the PHA cost. Starch is a biopolymer composed of glucose molecules, and regarded as one of the most abundant carbohydrates in nature. In terms of the molecular structure, starch is a macromolecular compound in which α-D-glucoside is connected by α-1,4-glycosidic linkages and α-1,6-glycosidic linkages (Hu et al., 2020). More than 70% of starch is used for the production of glucose, and amylase plays an important role in the process (Kmr et al., 2019). Amylase is a general term for enzymes with the ability to hydrolyze starch and glycogen, which could hydrolyze starch into low molecular structures (Chi et al., 2001; Zareian et al., 2010). Amylases are classified as α-amylase (EC 3.2.1.1) and β-amylase (EC 3.2.1.2) according to the type of sugar end groups generated during starch hydrolysis. The α-amylase, also known as α-1,4-glucose hydrolase, is more widely studied, with the ability of hydrolyzing starch into low molecular structures such as dextrin, maltose, and glucose (Chi et al., 2001). At present, many microorganisms have been employed to produce α-amylase, such as *Geobacillus* sp. (Mollania et al., 2010), *Bacillus subtilis* (Asgher et al., 2007), *B. licheniformis* (Shewale and Pandit, 2007), *B. stearothermophilus* (Chakraborty et al., 2000), and *Aspergillus oryzae* (Henriksen et al., 1999).

Previously, *V. alginolyticus* LHF01 isolated from a salt field was found to produce PHB from a variety of sugars and organic acids (Li et al., 2021). In this study, a genetic manipulation platform was developed for *V. alginolyticus* LHF01, and based on this, the utilization of starch by the strain and the overexpression of PHB synthesis operon and amylase were investigated. The direct utilization of starch by *V. alginolyticus* LHF01 would be helpful to reduce the cost of PHA production, making it as a promising bacterium for PHA production to achieve a breakthrough achievement in PHA commercialization.

2 Materials and methods

2.1 Strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* strain JM109 was used for molecular cloning and plasmid propagation, while *E. coli* strain S17-1 was used as the doner strain for conjugation (Simon et al., 1983). *V. alginolyticus* LHF01 with their engineered strains were employed for PHB production experiments.

The primers used in this study are listed in Supplementary Table S1. To overexpress the target genes in *V. alginolyticus* LHF01, pMCS1 was constructed by ligating the plasmid backbone fragment amplified from pBBR1MCS-2 (Kovach et al., 1995) with primers P1_F/R and the chloramphenicol fragment amplified from pSEVA341 (Tan et al., 2022) with primers cat_F/R by Gibson assembly. Plasmid pMCS1-phaBAPC was constructed to overexpress the PHB synthesis operon *phaBAPC* of *V. alginolyticus* LHF01. The *phaBAPC* and its native promoter were amplified from the genome of *V. alginolyticus* LHF01 using primers phaBAPC_F/R, and then inserted into the *Eco*RI/*Bam*HI sites of pMCS1 to generate pMCS1-phaBAPC. Similarly, pMCS1-03151, pMCS1-03416, and pMCS1-03713 were constructed to overexpress the amylase of *V. alginolyticus* LHF01. Primers amy2_F/R, amy3_F/R, and amy4_F/R were used to amplify the three different amylase gene and their own promoter, respectively. Then, the DNA fragment was inserted into the multiple cloning sites of pMCS1 by enzyme digestion and ligation. Plasmid pMCS1-03713-phaBAPC was constructed to co-overexpress *phaBAPC* and amylase gene 03713. The amylase gene fragment amplified from *V. alginolyticus* LHF01 using primers amy4_phaBAPC_F/R were digested with *Eco*RI/*Kpn*I and then cloned into the corresponding sites of pMCS1-phaBAPC to construct pMCS1-03713-phaBAPC.

2.2 Medium and culture conditions

During all experiments, *E. coli* was cultured in Luria-Bertani (LB) medium, while *V. alginolyticus* was cultured in TYS50 medium. TYS50 medium contained (g/L) NaCl 50, KCl 0.7, $CaCl_2{\cdot}2H_2O$ 1.4, $MgSO_4{\cdot}7H_2O$ 6.8, $MgCl_2{\cdot}6H_2O$ 5.4, $NaHCO_3$ 0.2, yeast extract 1, and peptone 5. When required, 1.5% (w/v) agar was added to obtain the corresponding solid medium. In the process of culturing *V. alginolyticus* LHF01 for PHB production, soluble starch was firstly used as carbon source, and the shake flask experiments were performed at 37°C and 200 rpm. The culture conditions were optimized in terms of pH, soluble starch concentration, and culture time. The pH was set at 5, 6, 6.5, 7, 7.5, 8, 8.5, 9, and 10, while the starch concentration gradient was 20, 30, and 40 g/L. The samples were taken every 2 hours to determine the cell growth, PHB accumulation, and starch consumption.

2.3 Genome sequencing

The genomic DNA of *V. alginolyticus* LHF01 was extracted using a Blood & Cell Culture DNA Midi Kit (Qiagen, United States) according to the manufacturer's protocol. DNA concentration and purity were determined *via* Qubit fluorometer and Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, United States).

TABLE 1 Strains and plasmids used in this study.

Name	Relevant characteristics	Reference
Strains		
E. coli JM109	*recA1 endA1 gyrA96 thi*-1 *hsdR17 supE44 relA1* Δ(*lac-proAB*)/F' [*traD36 proAB*$^{+}$ *lacI*^{q} *lacZ*ΔM15]	TaKaRa Bio Inc
E. coli S17-1	Donor strain in conjugation, harboring the *tra* genes of plasmid RP4 in the chromosome	Simon et al. (1983)
V. alginolyticus LHF01	Wild type, isolated from a salt field in China	This study
LHF01 (phaBAPC)	*V. alginolyticus* LHF01 harboring pMCS1-phaBAPC	This study
LHF01 (03151)	*V. alginolyticus* LHF01 harboring pMCS1-03151	This study
LHF01 (03416)	*V. alginolyticus* LHF01 harboring pMCS1-03416	This study
LHF01 (03713)	*V. alginolyticus* LHF01 harboring pMCS1-03713	This study
LHF01 (03713-phaBAPC)	*V. alginolyticus* LHF01 harboring pMCS1-03713-phaBAPC	This study
Plasmids		
pBBR1MCS-2	Broad range host plasmid, Kan^{R}	Kovach et al. (1995)
pSEVA341	Expression vector, Cat^{R}	Tan et al. (2022)
pMCS1	Broad range host plasmid, Cat^{R}	This study
pMCS1-phaBAPC	pMCS1 derived, carrying *phaBAPC* of LHF01	This study
pMCS1-03151	pMCS1 derived, carrying amylase gene 03151 of LHF01	This study
pMCS1-03416	pMCS1 derived, carrying amylase gene 03416 of LHF01	This study
pMCS1-03713	pMCS1 derived, carrying amylase gene 03713 of LHF01	This study
pMCS1-03713-phaBAPC	pMCS1 derived, carrying amylase gene 03713 and *phaBAPC* of LHF01	This study

DNA integrity was assessed by 1% agarose gel electrophoresis. Whole genome sequencing was performed on the MGISEQ-2000 platform and PacBio Sequel II system at BGI (Shenzhen, China).

2.4 Conjugation methods

The target plasmid was transferred into *V. alginolyticus* LHF01 by conjugation. In brief, the plasmid was firstly transformed into *E. coli* S17-1 to prepare the donor strain. *E. coli* S17-1 and *V. alginolyticus* strains were cultured in LB and TYS50 medium, respectively, with the relevant antibiotics to an OD_{600} of 1.5-2. Afterwards, 1.5 mL of donor and recipient cells were respectively placed in a centrifuge tube and centrifuged at 6,000 rpm for 5 min. The supernatant was discarded, and the cell pellets were washed once with 1 mL of the corresponding medium, and then suspended in 50 μL of TYS50 medium at a ratio of 1:1. The suspension was cultured for 4-8 h in a 37°C incubator. Finally, an inoculation loop was dipped into the bacterial solution, and then streaked on the TYS50 medium plate with appropriate antibiotics. The plate was cultured in a 37°C incubator for 36-48 h. A single clone was picked out, and put into the TYS50 medium with antibiotics for 24-48 h, and the bacterial solution was verified by 16S rDNA sequencing and stored for follow-up research.

2.5 Starch consumption and PHA production analysis

After treating with boiling water, the soluble starch can be dissolved and its solution is transparent at room temperature. Since starch is turned to blue when exposed with iodine, its consumption could be indicated by adding the iodine solution. To measure PHB accumulation, the strain culture was centrifuged at 10,000 g for 10 min, and washed twice with deionized water. Then the cell pellet was lyophilized to determine the cell dry weight (CDW) and PHB content. CDW was calculated based on the weight of the empty centrifuge tube, the weight of the centrifuge tube containing the cell pellet, and the volume of freeze-dried culture. The lyophilized cell was placed into an esterification tube, reacted with 2 mL of chloroform and 2 mL of esterification solution at 100°C for 4 h, and then shaken with 1 mL of deionized water, and set for stratification. The chloroform phase was taken for gas chromatography (GC) analysis as reported previously (Li et al., 2021).

2.6 Observation of cell morphology

In order to observe the cell morphological changes of *V. alginolyticus* strains when overexpressing the PHB synthesis operon, the wild type LHF01 and engineered recombinants were cultured in TYS50 medium with 20 g/L soluble starch for 18 h and observed by transmission electron microscopy (TEM). The shake flask culture was centrifuged to obtain the cell pellet, washed twice with pre-cooled phosphate buffered solution, slowly added with pre-cooled fixative solution, and then fixed for 12 h to obtain the cell samples. Then, the sample was sliced and observed according to the methods reported previously (Li et al., 2021).

3 Results and discussion

3.1 PHB production by *V. alginolyticus* using soluble starch

PHA is regarded as extremely competitive alternative to replace the petrochemical plastics, yet its high cost is always a major impediment. There have been numerous researches working on the exploitation of low-cost substrates and novel production hosts (Anjali et al., 2014; Vibhavee and Chanprateep, 2018). As an industrial raw material, starch has been applied for the production of high value-added products such as gas or liquid fuels, proteins, and sugars (Erdei et al., 2010; Lin et al., 2021; Li et al., 2022). However, many

10.3389/fbioe.2023.1130368

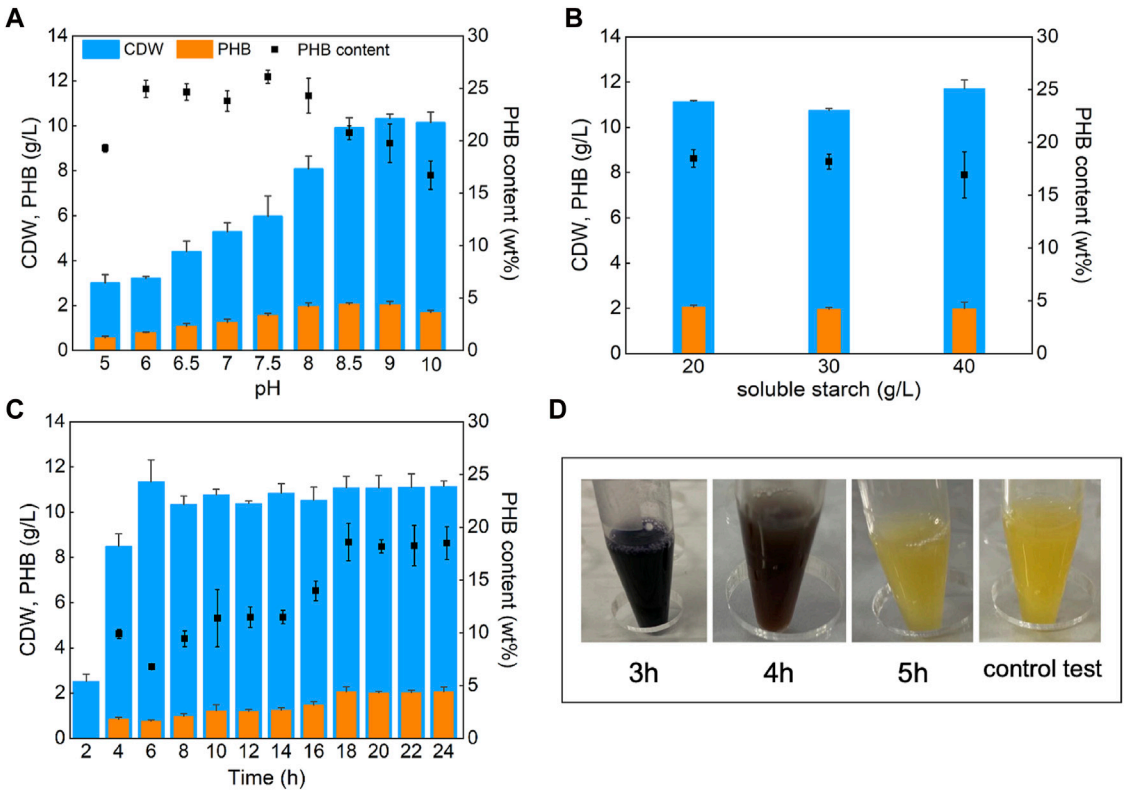


FIGURE 1
Effects of pH, substrate concentration, and culture time on PHB production. *V. alginolyticus* LHF01 was cultivated in TYS50 medium in shake flasks at 37°C, 200 rpm. Data are expressed as averages and standard deviations of three parallel experiments. **(A)**, the pH was set at 5, 6, 6.5, 7, 7.5, 8, 8.5, 9, and 10. **(B)**, the starch concentration gradient was 20, 30, and 40 g/L. **(C)**, the samples were taken every 2 hours to determine CDW and PHB accumulation. **(D)**, starch consumption was indicated by adding iodine solution to the cultures.

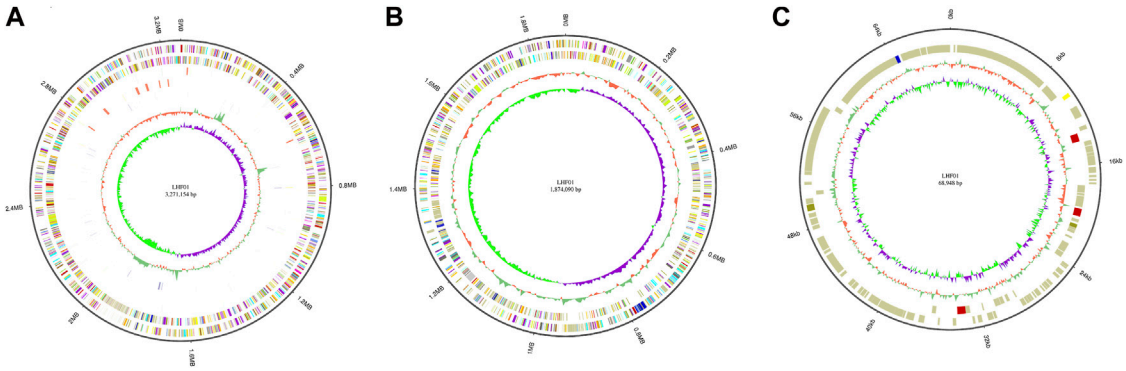


FIGURE 2
Genome map of *V. alginolyticus* LHF01. **(A)**, chromosome 1. **(B)**, chromosome 2. **(C)**, endogenous plasmid.

microorganisms are unable to utilize starch directly, and the cost of raw materials is increased due to the pretreatment of starch (Poomipuk et al., 2014). Interestingly, *Vibrio* strains have attracted increasing attentions due to their rapid growth rate and ability to assimilate starch (Hoff et al., 2020). *V. alginolyticus* was reported to accumulate PHB from a series of carbon sources (Li et al., 2021). The direct utilization of starch for PHA production and metabolic engineering were explored in *V. alginolyticus* LHF01 in this study.

V. alginolyticus LHF01 was firstly cultured with 20 g/L soluble starch as carbon source for 24 h to explore the effect of pH on PHB production (Figure 1A). Although the difference in pH had a certain effect on the CDW and PHB titers, *V. alginolyticus* LHF01 can

10.3389/fbioe.2023.1130368

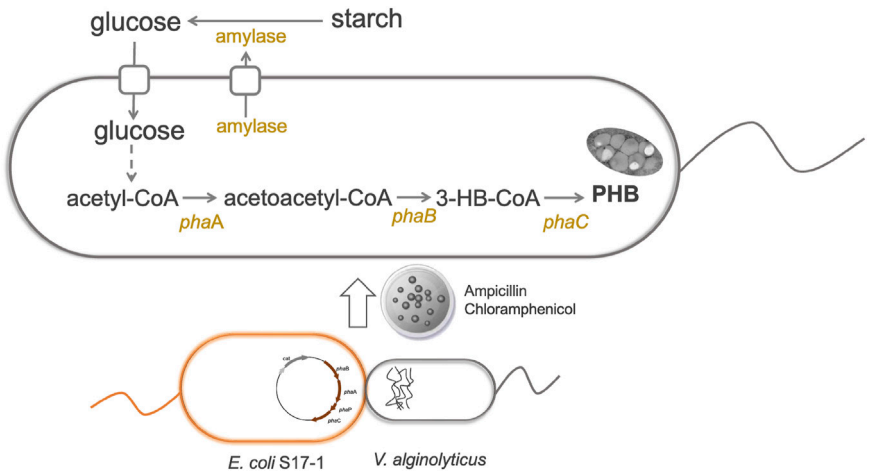


FIGURE 3
Construction of a genetic manipulation platform for improved PHB production in *V. alginolyticus* LHF01.

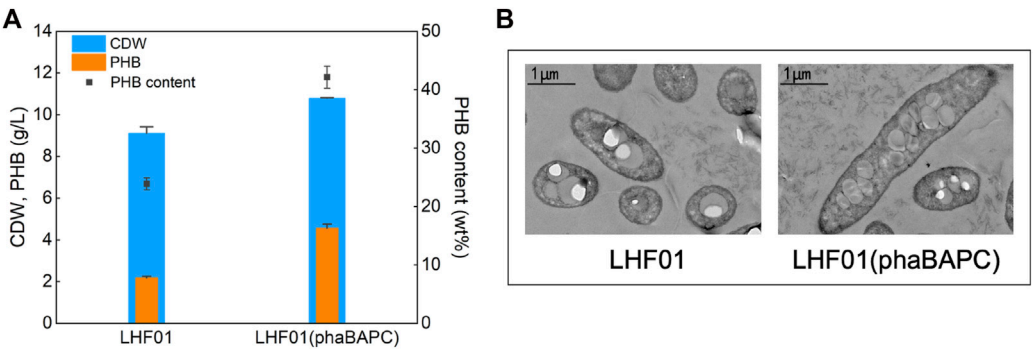


FIGURE 4
Effects of overexpressing PHB synthesis operon on PHB production.Strains were cultivated in TYS50 medium supplemented with 20 g/L soluble starch in shake flasks at 37°C and 200 rpm for 18 h **(A)**, CDW, PHB production, and PHB content in LHF01 and LHF01 (phaBAPC). **(B)**, TEM micrographs of LHF01 and LHF01 (phaBAPC).

synthesize PHB under a wide range of pH, indicating that it has strong viability and adaptable ability. PHB production by *V. alginolyticus* LHF01 showed an upward trend with the increase of pH value. When the pH was 8-10, CDW exceeded 8 g/L, and PHB titer was about 2 g/L, indicating that *V. alginolyticus* LHF01 is more suitable for survival under alkaline conditions. The highest titer PHB of 2.06 g/L was obtained at pH of 8.5. Therefore, the pH of the culture medium was maintained around 8.5 in the subsequent experiments of culturing *V. alginolyticus* LHF01.

Since the change of substrate concentration may also affect the growth of bacteria, the concentration of initial soluble starch was increased to 40 g/L to observe the change in PHB production. As shown in Figure 1B, the CDW and PHB titer did not change significantly with the increased soluble starch concentration. A possible reason was that the strain reached the growth threshold at 20 g/L of soluble starch, thus the increase of carbon concentration showed no assistance to further promote PHB accumulation. Then, *V. alginolyticus* LHF01 was cultured with a soluble starch concentration of 20 g/L for different time points to study the cell growth and PHB accumulation profiles. As shown in Figure 1C, the CDW of *V. alginolyticus* LHF01 reached 2.52 g/L after cultured for 2 h, up to 8.49 g/L for 4 h, and exceeded 10 g/L after 6 h, which confirmed the rapid cell growth of *V. alginolyticus* in shake flasks. Surprisingly, the PHB titer of *V. alginolyticus* LHF01 was only 0.84 g/L at 4 h of cultivation, indicating that PHB was synthesized during the late stage of cell growth. After 18 h of cultivation, PHB titer reached the highest level. It is well known that starch or soluble starch is turned to blue when exposed to iodine. To explore the consumption rate of soluble starch, iodine was used to determine whether soluble starch was completely consumed. As shown in Figure 1D, soluble starch was exhausted by *V. alginolyticus* LHF01 after 5 h of cultivation, indicating that the strain has superior ability to degrade soluble starch. The rapid cell growth and efficient utilization of starch make *V. alginolyticus* a promising candidate for microbial production of building block chemicals or polymers using starch as the substrate.

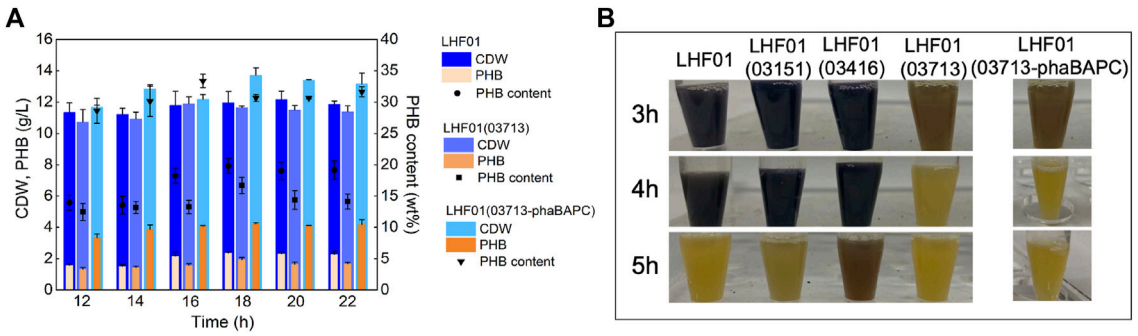


FIGURE 5
Effects of co-overexpressing amylase and PHB synthesis operon on PHB production.Strains were cultivated in TYS50 medium supplemented with 20 g/L soluble starch in shake flasks at 37°C and 200 rpm. **(A)**, CDW, PHB production, and PHB content in LHF01, LHF01 (03713), and LHF01 (03713-phaBAPC). **(B)**, starch consumption was indicated by adding iodine solution to the cultures.

3.2 Complete genome sequencing of *V. alginolyticus* LHF01

Whole genome sequencing has the highest resolution in bacterial genetic research, and is widely used in microbial traceability, transmission, and population structure identification. Therefore, the genomic characterization of *V. alginolyticus* LHF01 was performed. Its complete genome was sequenced and reassembled from Celera Assembler 8.3 (Denisov et al., 2013). The complete genome sequence was deposited in GenBank database under the accession numbers of CP087876.1, CP087877.1, and CP087878.1. Results indicated that the genome of *V. alginolyticus* LHF01 contains two chromosomes and one plasmid (Figure 2). The full length of chromosome one is 3,271,154 bp, with GC content of 44.74%, and the full length of chromosome two is 1,874,090 bp, with GC content of 44.68%. The full length of plasmid DNA was 68,948 bp and GC content was 42.24%. The annotation was carried out with the National Centre for Biotechnology Information Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) Version 4.9 (https://www.ncbi.nlm.nih.gov/genome/annotation_prok). From the 4,882 predicted genes identified, a total of 4,639 protein-coding sequences (CDSs) were determined, and 128 tRNA genes, 37 complete rRNA genes (13 for 5S, and 12 each for 16S and 23S), and four ncRNAs were annotated.

In the metabolic pathways of starch degradation and PHB production, amylase and PHB synthesis-related enzymes are indispensable. There are multiple genes encoding amylase, and the gene locus tag number is 03151, 03416, and 03713, respectively. A phylogenetic tree of the amylases of *V. alginolyticus* LHF01 was constructed, and results indicated that they are all α-amylase and belong to the glycoside hydrolase family 13 (Kuriki et al., 2006) (Supplementary Figure S1).

Previous studies reported that there were genes encoding PHA synthesis enzymes existed in the genome of *Vibrio* strains including *V. harvinii* (Mohandas et al., 2016), *V. natriureus* (Chien et al., 2007), and *V. azureus* (Sasidharan et al., 2014). Genome analysis revealed the presence of PHB synthesis operon *phaBAPC* in *V. alginolyticus* LHF01. The operon comprised of four genes including *phaB*, *phaA*, *phaP*, and *phaC*, encoding acetoacetyl-CoA reductase, acetyl-CoA acetyltransferase, phasin, and PHA synthase, respectively. The amino acid sequence of PHA synthase in *V. alginolyticus* LHF01 belongs to Class I PHA synthase. The genes responsible for PHA biosynthesis usually clustered together in the bacterial genome (Rehm and Steinbüchel, 1999), such as the most well-known genetic organization of *phaCAB* operon in *Ralstonia eutropha* H16 (Peoples and Sinskey, 1989), the *phaCA* operon in *Jeongeupia* sp. USM3 (Zain et al., 2020), the *phaEC* operon in *Neptunomonas concharum* (Pu et al., 2020b). Interestingly, the presence of phasin gene in a PHA operon is rarely reported. The neighborhood genetic organization of *phaP* and *phaA*, *phaB*, *phaC* genes may play a unique role in the regulation of PHB metabolism in *V. alginolyticus*, which deserves further study. The genome sequencing of *V. alginolyticus* would provide a good foundation for further in-depth metabolic engineering for improved substrate utilization and PHB production.

3.3 Effects of overexpressing PHB synthesis operon on PHB production

The genetic modification of host cells is one of the most powerful tools in engineering bacteria for microbial fermentation. To begin with, molecular biology research requires introducing foreign plasmids into the target hosts. Generally, developing DNA transformation methods for non-model organism poses unknown challenges. Fortunately, the broad host pBBR1 origin was reported to be transformable into *V. natriegens via* conjugation (Tschirhart et al., 2019). No chloramphenicol resistance was observed for *V. alginolyticus* (Li et al., 2021), thus plasmid pMCS1 harboring pBBR1 origin and chloramphenicol resistance gene was constructed to study its possible application to transfer foreign DNA into *V. alginolyticus* LHF01. We firstly tried electroporation and chemical transformation using pMCS1 as the shuttle vector. However, it was found that both methods failed and no transformant was observed despite of many attempts. Recently, the deletion of gene clusters involved in the biosynthesis of exopolysaccharides and O-antigen was proved to improve the permeability of exogenous DNA into cells and enabled the electrotransformation in *Halomonas bluephagenesis* (Xu et al., 2022). The electroporation method for *V. alginolyticus* LHF01 was worthy of further investigation. Next, a DNA transformation procedure through conjugation was established and pMCS1 was successfully introduced into *V. alginolyticus* LHF01 by using *E. coli* S17-1 as the assistant strain (Figure 3). Overexpression of

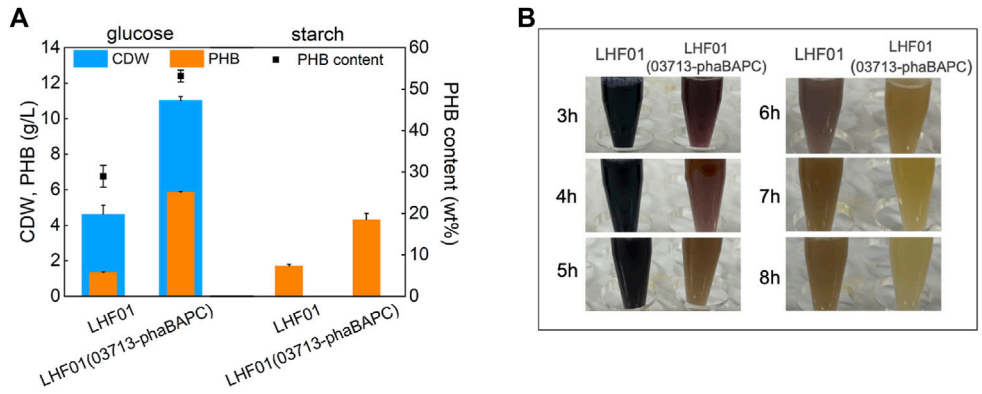


FIGURE 6
PHB production by *V. alginolyticus* strains using corn starch or glucose.Strains were cultivated in TYS50 medium supplemented with 20 g/L insoluble corn starch or glucose in shake flasks at 37°C and 200 rpm. **(A)**, CDW, PHB production, and PHB content in LHF01 and LHF01 (03713-phaBAPC). **(B)**, starch consumption was indicated by adding iodine solution to the cultures.

TABLE 2 PHA production by microorganisms using starch as the carbon source.

Strain	Carbon source	Products	CDW (g/L)	PHA content (wt%)	PHA (g/L)	Reference
Aeromonas sp. KC007-R1	Starch	PHB	1.83	32.7	0.6	Chien and Ho (2008)
Bacillus cereus CFR06	Soluble starch	PHB	2.14	46	1.0	Halami (2008)
Massilia sp. UMI-21	Corn starch	PHB	3.95	30.3	1.20	Han et al. (2014)
Halogeometricum borinquense	Soluble starch	PHBV	6.2	74.19	4.6	Salgaonkar et al. (2019)
Halorubrum chaoviator CEJ34-14	Soluble starch	PHB	0.22	9.25	0.02	Karray et al. (2021)
Natrinema pallidum CEJ5-14	Soluble starch	PHB	0.56	7.11	0.04	Karray et al. (2021)
Halolamina sp. NRS_35	Starch	PHB	-	-	0.042	Hagagy et al. (2022)
Halolamina sp. NRS_38	Starch	PHB	-	-	0.037	Hagagy et al. (2022)
E. coli SKB99	Starch	PHB	2.16	57.4	1.24	Bhatia et al. (2015)
Halomonas bluephagenesis TN04	Corn starch	PHB	9.5	51.5	4.89	Lin et al. (2021)
V. alginolyticus	Soluble starch	PHB	10.80	42.13	4.55	This study
V. alginolyticus	Corn starch	PHB	-	-	4.32	This study

the endogenous PHA synthesis genes is a common metabolic strategy for directing the carbon metabolic flux towards PHA synthesis pathway and strengthening PHA production (Khetkorn et al., 2016; Zhao et al., 2019; Tang et al., 2022). For example, the application of promoter engineering to enhance transcription of PHA synthase gene increased PHA accumulation in *Pseudomonas mendocina* (Zhao et al., 2019). To explore the role of native PHB synthesis genes on the production of PHB, *phaBAPC* operon was cloned with the native promoter and ligated into pMCS1 to construct the plasmid pMCS1-phaBAPC. The constructed plasmid was transferred into *V. alginolyticus* LHF01 by conjugation to obtain the recombinants with ability to overexpress the PHB synthesis operon. The recombinant strain LHF01 (phaBAPC) was verified by 16S rDNA sequencing.

Soluble starch was used to cultivate the recombinant strain harboring extra *phaBAPC* operon and the wild type control. As shown in Figure 4A, PHB titer of 4.55 g/L and PHB content was 42.13% were obtained in strain LHF01 (phaBAPC). The PHB titer was increased by more than 100% compared with the control strain LHF01. Therefore, the overexpression of *phaBAPC* can well promote the conversion of soluble starch to PHB, and it also demonstrates the success of genetic manipulation in *V. alginolyticus*. In addition, the cells of *V. alginolyticus* LHF01 and LHF01 (phaBAPC) grown on soluble starch were observed by TEM. It was obvious that electron translucent inclusions can be observed in both bacterial cells, yet LHF01 (phaBAPC) has significantly more intracellular particles than *V. alginolyticus* LHF01 (Figure 4B). Both PHA synthase and phasins existed on the surface of PHA granules by covalent and hydrophobic interactions (Cai et al., 2011). It was proposed that the phasins affected the activity of PHA synthase and determined the number and size of PHA granules (Mezzina and Pettinari, 2016). Herein, similar phenomenon was also

observed in *V. alginolyticus*, and the overexpression of *phaBAPC* is beneficial to the PHB production.

3.4 Effects of overexpressing amylase on PHB production

Next, the three amylase genes of *V. alginolyticus* LHF01 were cloned to construct plasmids pMCS1-03151, pMCS1-03416, and pMCS1-03713, respectively. Recombinants LHF01 (03151), LHF01 (03416), and LHF01 (03713) were obtained by conjugation. Shake flask experiments were performed to study the consumption rate of soluble starch. As shown in Figure 5B, the samples of LHF01 (03151), LHF01 (03416), and control strain LHF01 were no longer turned blue at about 5 h, while the sample of LHF01 (03713) was turned yellow at 3-4 h. Among the three amylases, the overexpression of amylase 03713 significantly increased the consumption of soluble starch. However, the overexpression of amylase did not effectively improve PHB accumulation (Supplementary Figure S2).

Afterwards, strain LHF01 (03713-phaBAPC) was constructed to explore the synergistic effect of PHB synthesis operon and amylase on substrate consumption and PHB accumulation. The PHB titer of *V. alginolyticus* LHF01, LHF01 (03713) and LHF01 (03713-phaBAPC) all reached the maximum at 18 h, but it was obvious that LHF01 (03713-phaBAPC) had the PHB highest titer of 4.20 g/L at this time, and the PHB content reached 30.66% (Figure 5A). Furthermore, the culture of LHF01 (03713-phaBAPC) was turned to yellow at 3-4 h, indicating that LHF01 (03713-phaBAPC) exhibited faster consumption rate of soluble starch compared with *V. alginolyticus* LHF01 (Figure 5B). It can be said that these phenomena indicated that the co-expression of 03713 gene with *phaBAPC* not only increased the rate of soluble starch consumption, but also increased the titer of PHB.

It had to be admitted that the cost of soluble starch is relatively high due to the reliance on chemical or other means to denature starch. Therefore, soluble starch was replaced with insoluble corn starch to explore the direct utilization of starch by *V. alginolyticus*. Strain LHF01 and LHF01 (03713-phaBAPC) were cultivated in TYS50 medium supplemented with 20 g/L corn starch for 18 h, and starch consumption and PHB titer were recorded. Glucose was also employed as the control. Due to the insolubility of corn starch, the accurate CDW cannot be measured. The PHB titer of wide type strain LHF01 was 1.70 g/L, a litter higher than that of 1.33 g/L when glucose was used as the carbon source. In terms of corn starch consumption, starch was not completely consumed after 8 h cultivation by LHF01, showing that the utilization ability of corn starch was weaker than that of soluble starch. It is worth mentioning that, the PHB titer of LHF01 (03713-phaBAPC) reached 4.32 g/L with corn starch, slightly lower than that of 5.85 g/L using glucose as carbon source (Figure 6A). In addition, the sample of LHF01 (03713-phaBAPC) was turned into yellow at 6-7 h, indicating that the starch consumption ability of LHF01 (03713-phaBAPC) was much higher than the control strain LHF01 (Figure 6B).

Several strains have been reported to utilize starch for PHA production, including *Aeromonas*, *Bacillus cereus*, *Massilia*, *Halogeometricum borinquense*, and *Halolamina* (Table 2). Generally, the PHA titers and PHA content in the isolated wild type strains were not high. To improve starch utilization and PHA accumulation, metabolic engineering strategies have been applied in *E. coli* and *H. bluephagenesis*, which cannot assimilate starch naturally. The recombinant *E. coli* carrying amylase gene from *Panibacillus* sp. and PHB synthesis genes from *R. eutropha* accumulated 1.24 g/L of PHB using starch (Bhatia et al., 2015). Engineered *H. bluephagenesis* harboring an amylase with suitable signal peptide was able to grow on corn starch to 9.5 g/L cell dry weight containing 51.5% PHB (Lin et al., 2021). Compared to previous reported studies, *V. alginolyticus* LHF01 indicated the fastest starch consumption rate. For example, the engineered *V. alginolyticus* could exhaust 20 g/L corn starch in 7 h and accumulate 4.32 g/L PHB after 18 h cultivation, while the shake flask experiments for PHA production by other reported strains using starch were usually performed for 48 h or even longer. These results demonstrated *V. alginolyticus* to be a superior chassis for production of PHB from low-cost corn starch. Further development of genetic manipulation techniques and rational engineering will help to improve PHB production from starch by *V. alginolyticus*."

4 Conclusion

In this study, a genetic manipulation platform based on conjugation was established for the first-time in *V. alginolyticus* LHF01, and based on this, the utilization of starch by the strain and the overexpression of PHB synthesis operon and amylase were investigated. *V. alginolyticus* LHF01 was able to consume 20 g/L soluble starch in 5 h shake flask cultivation, achieving PHB titer of 2.06 g/L. Afterwards, the genome of *V. alginolyticus* LHF01 was sequenced and analyzed. With the overexpression of its *phaBAPC* operon, PHB titer was improved to 4.55 g/L, and the consumption of soluble starch was increased with the overexpression of its native amylase. Notably, the engineered *V. alginolyticus* harboring extra amylase and PHB synthesis operon was successfully used to degrade 20 g/L corn starch in 7 h shake flask cultivation, with PHB titer of 4.32 g/L. The use of *V. alginolyticus* points to the possibility of rapid cell growth for production of chemicals and biopolymers using starch, a sustainable resource with cost much lower than glucose. Further engineering of *V. alginolyticus* may open new applications in metabolic engineering and synthetic biology research.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, CP087876.1. https://www.ncbi.nlm.nih.gov/genbank/, CP087877.1. https://www.ncbi.nlm.nih.gov/genbank/, CP087878.1.

Author contributions

H-FL: Conceptualization, Data curation, Formal analysis, Experimental design and setup, Writing—original draft. LT: Investigation, Experimental design and setup. L-HF: Management and Supervision, Writing—review and editing, Funding acquisition.

Z-JL: Conceptualization, Data curation, Management and Supervision, Writing—review and editing, Funding acquisition.

Funding

This study was supported by grants from the National Key R&D Program of China (2021YFC2103100, 2018YFA0900200), and Quanzhou City Science & Technology Program of China (No. 2021CT006).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1130368/full#supplementary-material

References

Anjali, M., Sukumar, C., Kanakalakshmi, A., and Shanthi, K. (2014). Enhancement of growth and production of polyhydroxyalkanoates by *Bacillus subtilis* from agro-industrial waste as carbon substrates. *Compos Interface* 21, 111–119. doi:10.1080/15685543.2013.834200

Asgher, M., Bhatti, H. N., Shah, S. A. H., Asad, M. J., and Legge, R. L. (2007). Decolorization potential of mixed microbial consortia for reactive and disperse textile dyestuffs. *Biodegradation* 18, 311–316. doi:10.1007/s10532-006-9065-7

Bhatia, S. K., Shim, Y. H., Jeon, J. M., Brigham, C. J., Kim, Y. H., Kim, H. J., et al. (2015). Starch based polyhydroxybutyrate production in engineered *Escherichia coli*. *Bioprocess Biosyst. Eng.* 38, 1479–1484. doi:10.1007/s00449-015-1390-y

Cai, L., Tan, D., Aibaidula, G., Dong, X. R., Chen, J. C., Tian, W. D., et al. (2011). Comparative genomics study of polyhydroxyalkanoates (PHA) and ectoine relevant genes from *Halomonas* sp. TD01 revealed extensive horizontal gene transfer events and co-evolutionary relationships. *Microb. Cell. Fact.* 10, 88. doi:10.1186/1475-2859-10-88

Chakraborty, K., Bhattacharyya, B. K., and Sen, S. K. (2000). Purification and characterization of a thermostable alpha-amylase from *Bacillus stearothermophilus*. *Folia Microbiol.* 45, 207–210. doi:10.1007/bf02908945

Chi, Z. M., Liu, J., and Zhang, W. (2001). Trehalose accumulation from soluble starch by *Saccharomycopsis fibuligera sdu*. *Enzyme Microb. Technol.* 28, 240–245. doi:10.1016/s0141-0229(00)00318-5

Chien, C. C., Chen, C. C., Choi, M. H., Kung, S. S., and Wei, Y. H. (2007). Production of poly-β-hydroxybutyrate (PHB) by *Vibrio* spp. isolated from marine environment. *J. Biotechnol.* 132, 259–263. doi:10.1016/j.jbiotec.2007.03.002

Chien, C. C., and Ho, L. Y. (2008). Polyhydroxyalkanoates production from carbohydrates by a genetic recombinant *Aeromonas* sp. *Lett. Appl. Microbiol.* 47, 587–593. doi:10.1111/j.1472-765X.2008.02471.x

Choi, S. Y., Cho, I. J., Lee, Y., Kim, Y. J., Kim, K. J., and Lee, S. Y. (2020). Microbial polyhydroxyalkanoates and nonnatural polyesters. *Adv. Mater* 32, e1907138. doi:10.1002/adma.201907138

Cui, Y. W., Zhang, H. Y., Lu, P. F., and Peng, Y. Z. (2016). Effects of carbon sources on the enrichment of halophilic polyhydroxyalkanoate-storing mixed microbial culture in an aerobic dynamic feeding process. *Sci. Rep.* 6, 30766. doi:10.1038/srep30766

Dahman, Y., and Ugwu, C. U. (2014). Production of green biodegradable plastics of poly(3-hydroxybutyrate) from renewable resources of agricultural residues. *Bioproc Biosyst. Eng.* 37, 1561–1568. doi:10.1007/s00449-014-1128-2

Dalia, T. N., Hayes, C. A., Stolyar, S., Marx, C. J., McKinlay, J. B., and Dalia, A. B. (2017). Multiplex genome editing by natural transformation (MuGENT) for synthetic biology in *Vibrio natriegens*. *Acs Syn. Biol.* 6, 1650–1655. doi:10.1021/acssynbio.7b00116

Denisov, G., Walenz, B., Halpern, A. L., Miller, J., Axelrod, N., Levy, S., et al. (2013). Consensus generation and variant detection by Celera Assembler. *Bioinformatics* 24, 1035–1040. doi:10.1093/bioinformatics/btn074

Erdei, B., Barta, Z., Sipos, B., Reczey, K., Galbe, M., and Zacchi, G. (2010). Ethanol production from mixtures of wheat straw and wheat meal. *Biotechnol. Biofuels* 3, 16. doi:10.1186/1754-6834-3-16

HaasJin, R. B., and Zepf, F. T. (2008). Production of poly(3-hydroxybutyrate) from waste potato starch. *Biosci. Biotech. Bioch* 72, 253–256. doi:10.1271/bbb.70503

Hagagy, N., Saddiq, A. A., Tag, H. M., Selim, S., AbdElgawad, H., and Martínez-Espinosa, R. M. (2022). Characterization of polyhydroxybutyrate, PHB, synthesized by newly isolated haloarchaea *Halolamina* spp. *Molecules* 27, 7366. doi:10.3390/molecules27217366

Halami, P. M. (2008). Production of polyhydroxyalkanoate from starch by the native isolate *Bacillus cereus* CFR06. *World J. Microbiol. Biotechnol.* 24, 805–812. doi:10.1007/s11274-007-9543-z

Han, X., Satoh, Y., Kuriki, Y., Seino, T., Fujita, S., Suda, T., et al. (2014). Polyhydroxyalkanoate production by a novel bacterium *Massilia* sp. UMI-21 isolated from seaweed, and molecular cloning of its polyhydroxyalkanoate synthase gene. *J. Biosci. Bioeng.* 118, 514–519. doi:10.1016/j.jbiosc.2014.04.022

Henriksen, A. L. S., Carlsen, M., de Bang, H., and Nielsen, J. (1999). Kinetics of alpha-amylase secretion in *Aspergillus oryzae*. *Biotechnol. Bioeng.* 65 (1), 76–82. doi:10.1002/(sici)1097-0290(19991005)65:1<76::aid-bit9>3.0.co;2-d

Hoff, J., Daniel, B., Stukenberg, D., Thuronyi, B. W., Waldminghaus, T., and Fritz, G. (2020). *Vibrio natriegens*: An ultrafast-growing marine bacterium as emerging synthetic biology chassis. *Environ. Microbiol.* 22, 4394–4408. doi:10.1111/1462-2920.15128

Hong, J. W., Song, H. S., Moon, Y. M., Hong, Y. G., Bhatia, S. K., Jung, H. R., et al. (2019). Polyhydroxybutyrate production in halophilic marine bacteria *Vibrio proteolyticus* isolated from the Korean peninsula. *Bioproc Biosyst. Eng.* 42, 603–610. doi:10.1007/s00449-018-02066-6

Hu, B., Li, Y., Zhu, S., Zhang, H., Jing, Y., Jiang, D., et al. (2020). Evaluation of biohydrogen yield potential and electron balance in the photo-fermentation process with different initial pH from starch agricultural leftover. *Bioresour. Technol.* 305, 122900. doi:10.1016/j.biortech.2020.122900

Karray, F., Ben Abdallah, M., Baccar, N., Zaghden, H., and Sayadi, S. (2021). Production of poly(3-Hydroxybutyrate) by Haloarcula, Halorubrum, and Natrinema haloarchaeal genera using starch as a carbon source. *Archaea* 2021, 1–10. doi:10.1155/2021/8888712

Khetkorn, W., Incharoensakdi, A., Lindblad, P., and Jantaro, S. (2016). Enhancement of poly-3-hydroxybutyrate production in *Synechocystis* sp. PCC 6803 by overexpression of its native biosynthetic genes. *Bioresour. Technol.* 214, 761–768. doi:10.1016/j.biortech.2016.05.014

Kmr, K., Husaini, A., Sing, N. N., Tasnim, T., Mohd, S. F., Hussain, H., et al. (2019). Characterization and expression in *Pichia pastoris* of a raw starch degrading glucoamylase (GA2) derived from *Aspergillus flavus* NSH9. *Protein Express Purif.* 164, 105462. doi:10.1016/j.pep.2019.105462

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. doi:10.1016/0378-1119(95)00584-1

Kuriki, T., Takata, H., Yanase, M., Ohdan, K., Fujii, K., Terada, Y., et al. (2006). The concept of the.ALPHA. Amylase family: A rational tool for interconverting glucanohydrolases/glucanotransferases, and their specificities. *J. Appl. Glycosci.* 53, 155–161. doi:10.5458/jag.53.155

Lee, H. H., Ostrov, N., Wong, B. G., Gold, M. A., Khalil, A. S., and Church, G. M. (2019). Functional genomics of the rapidly replicating bacterium *Vibrio natriegens* by CRISPRi. *Nat. Microbiol.* 4, 1105–1113. doi:10.1038/s41564-019-0423-8

Lee, H. S., Lee, H. J., Kim, S. H., Cho, J. Y., Suh, M. J., Ham, S., et al. (2021). Novel phasins from the *Arctic Pseudomonas* sp. B14-6 enhance the production of polyhydroxybutyrate and increase inhibitor tolerance. *Int. J. Biol. Macromol.* 190, 722–729. doi:10.1016/j.ijbiomac.2021.08.236

Li, C. L., Ruan, H. Z., Liu, L. M., Zhang, W. G., and Xu, J. Z. (2022). Rational reformation of *Corynebacterium glutamicum* for producing L-lysine by one-step fermentation from raw corn starch. *Appl. Microbiol. Biotechnol.* 106, 145–160. doi:10.1007/s00253-021-11714-z

Li, H. F., Wang, M. R., Tian, L. Y., and Li, Z. J. (2021). Production of polyhydroxyalkanoates (PHAs) by *Vibrio alginolyticus* strains isolated from salt fields. *Molecules* 26, 6283. doi:10.3390/molecules26206283

Lin, Y., Guan, Y., Dong, X., Ma, Y., Wang, X., Leng, Y., et al. (2021). Engineering *Halomonas bluephagenesis* as a chassis for bioproduction from starch. *Metab. Eng.* 64, 134–145. doi:10.1016/j.ymben.2021.01.014

Liu, H., Kumar, V., Jia, L., Sarsaiya, S., Kumar, D., Juneja, A., et al. (2021). Biopolymer poly-hydroxyalkanoates (PHA) production from apple industrial waste residues: A review. *Chemosphere* 284, 131427. doi:10.1016/j.chemosphere.2021.131427

Meng, D. C., Shen, R., Yao, H., Chen, J. C., Wu, Q., and Chen, G. Q. (2014). Engineering the diversity of polyesters. *Curr. Opin. Biotechnol.* 29, 24–33. doi:10.1016/j.copbio.2014.02.013

Mezzina, M. P., and Pettinari, M. J. (2016). Phasins, multifaceted polyhydroxyalkanoate granule-associated proteins. *Appl. Environ. Microbiol.* 82, 5060–5067. doi:10.1128/AEM.01161-16

Mohandas, S. P., Balan, L., Lekshmi, N., Cubelio, S. S., Philip, R., and Bright Singh, I. S. (2016). Production and characterization of polyhydroxybutyrate from *Vibrio harveyi* MCCB 284 utilizing glycerol as carbon source. *J. Appl. Microbiol.* 122, 698–707. doi:10.1111/jam.13359

Mollania, N., Khajeh, K., Hosseinkhani, S., and Dabirmanesh, B. (2010). Purification and characterization of a thermostable phytate resistant alpha-amylase from *Geobacillus* sp LH8. *Int. J. Biol. Macromol.* 46, 27–36. doi:10.1016/j.ijbiomac.2009.10.010

Pagliano, G., Galletti, P., Samori, C., Zaghini, A., and Torri, C. (2021). Recovery of polyhydroxyalkanoates from single and mixed microbial cultures: A review. *Front. Bioeng. Biotechnol.* 9, 624021. doi:10.3389/fbioe.2021.624021

Peoples, O. P., and Sinskey, A. J. (1989). Poly-beta-hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus* H16: Identification and characterization of the PHB polymerase gene (*phbC*). *J. Biol. Chem.* 264, 15298–15303. doi:10.1016/S0021-9258(19)84825-1

Poomipuk, N., Reungsang, A., and Plangklang, P. (2014). Poly-beta-hydroxyalkanoates production from cassava starch hydrolysate by *Cupriavidus* sp. KKU38. *Int. J. Biol. Macromol.* 65, 51–64. doi:10.1016/j.ijbiomac.2014.01.002

Pu, N., Hu, P., Shi, L. L., and Li, Z. J. (2020a). Microbial production of poly(3-hydroxybutyrate) from volatile fatty acids using the marine bacterium *Neptunomonas concharum*. *Bioresour. Technol. Rep.* 11, 100439. doi:10.1016/j.biteb.2020.100439

Pu, N., Wang, M. R., and Li, Z. J. (2020b). Characterization of polyhydroxyalkanoate synthases from the marine bacterium *Neptunomonas concharum* JCM17730. *J. Biotechnol.* 319, 69–73. doi:10.1016/j.jbiotec.2020.06.002

Rehm, B. H. A., and Steinbüchel, A. (1999). Biochemical and genetic analysis of PHA synthases and other proteins required for PHA synthesis. *Int. J. Biol. Macromol.* 25, 3–19. doi:10.1016/s0141-8130(99)00010-0

Roh, H., Yun, E. J., Lee, S., Ko, H. J., Kim, S., Kim, B. Y., et al. (2012). Genome sequence of *Vibrio* sp strain EJY3, an agarolytic marine bacterium metabolizing 3,6-anhydro-L-galactose as a sole carbon source. *J. Bacteriol.* 194, 2773–2774. doi:10.1128/jb.00303-12

Sagong, H. Y., Son, H. F., Choi, S. Y., Lee, S. Y., and Kim, K. J. (2018). Structural insights into polyhydroxyalkanoates biosynthesis. *Trends Biochem. Sci.* 43, 790–805. doi:10.1016/j.tibs.2018.08.005

Salgaonkar, B. B., Mani, K., and Bragança, J. M. (2019). Sustainable bioconversion of cassava waste to poly(3-hydroxybutyrate-co-3-hydroxyvalerate) by *Halogeometricum borinquense* strain E3. *J. Polym. Environ.* 27, 299–308. doi:10.1007/s10924-018-1346-9

Sasidharan, R. S., Bhat, S. G., and Chandrasekaran, M. (2014). Biocompatible polyhydroxybutyrate (PHB) production by marine *Vibrio azureus* BTKB33 under submerged fermentation. *Ann. Microbiol.* 65, 455–465. doi:10.1007/s13213-014-0878-z

Shewale, S. D., and Pandit, A. B. (2007). Hydrolysis of soluble starch using *Bacillus licheniformis* alpha-amylase immobilized on superporous CELBEADS. *Carbohyd Res.* 342, 997–1008. doi:10.1016/j.carres.2007.02.027

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. *Nat. Biotech.* 1, 784–791. doi:10.1038/nbt1183-784

Tan, B., Zheng, Y., Yan, H., Liu, Y., and Li, Z. J. (2022). Metabolic engineering of *Halomonas bluephagenesis* to metabolize xylose for poly-3-hydroxybutyrate production. *Biochem. Eng. J.* 187, 108623. doi:10.1016/J.BEJ.2022.108623

Tang, R., Peng, X., Weng, C., and Han, Y. (2022). The overexpression of phasin and regulator genes promoting the synthesis of polyhydroxybutyrate in *Cupriavidus necator* H16 under nonstress conditions. *Appl. Environ. Microbiol.* 88, e0145821. doi:10.1128/AEM.01458-21

Tian, L., Li, H., Song, X., Ma, L., and Li, Z. J. (2022). Production of polyhydroxyalkanoates by a novel strain of *Photobacterium* using soybean oil and corn starch. *J. Environ. Chem. Eng.* 10, 108342. doi:10.1016/j.jece.2022.108342

Tschirhart, T., Shukla, V., Kelly, E. E., Schultzhaus, Z., NewRingeisen, E., Erickson, J. S., et al. (2019). Synthetic biology tools for the fast-growing marine bacterium *Vibrio natriegens*. *ACS Synth. Biol.* 8, 2069–2079. doi:10.1021/acssynbio.9b00176

Vibhavee, S., and Chanprateep, N. S. (2018). Use of agro-industrial residue from the canned pineapple industry for polyhydroxybutyrate production by *Cupriavidus necator* strain A-04. *Biotechnol. Biofuels* 11, 202. doi:10.1186/s13068-018-1207-8

Wang, M. R., Li, H. F., Yi, J. J., Tao, S. Y., and Li, Z. J. (2022). Production of polyhydroxyalkanoates by three novel species of *Marinobacterium*. *Int. J. Biol. Macromol.* 195, 255–263. doi:10.1016/j.ijbiomac.2021.12.019

Weinstock, M. T., Hesek, E. D., Wilson, C. M., and Gibson, D. G. (2016). *Vibrio natriegens* as a fast-growing host for molecular biology. *Nat. Methods* 13, 849–851. doi:10.1038/nmeth.3970

Xu, T., Chen, J., Mitra, R., Lin, L., Xie, Z., Chen, G. Q., et al. (2022). Deficiency of exopolysaccharides and O-antigen makes *Halomonas bluephagenesis* self-flocculating and amenable to electrotransformation. *Commun. Biol.* 5, 623–711. doi:10.1038/s42003-022-03570-y

Zain, N. A. A., Ng, L. M., Foong, C. P., Tai, Y. T., Nanthini, J., and Sudesh, K. (2020). Complete genome sequence of a novel polyhydroxyalkanoate (PHA) producer, *Jeongeupia* sp. USM3 (JCM 19920) and characterization of its PHA synthases. *Curr. Microbiol.* 77, 500–508. doi:10.1007/s00284-019-01852-z

Zareian, S., Khajeh, K., Ranjbar, B., Dabirmanesh, B., Ghollasi, M., and Mollania, N. (2010). Purification and characterization of a novel amylopullulanase that converts pullulan to glucose, maltose, and maltotriose and starch to glucose and maltose. *Enzyme Microb. Technol.* 46, 57–63. doi:10.1016/j.enzmictec.2009.09.012

Zhao, F., Liu, X., Kong, A., Zhao, Y., Fan, X., Ma, T., et al. (2019). Screening of endogenous strong promoters for enhanced production of medium-chain-length polyhydroxyalkanoates in *Pseudomonas mendocina* NK-01. *Sci. Rep.* 9, 1798–1813. doi:10.1038/s41598-019-39321-z

Zhao, J., Cui, Y. W., Zhang, H. Y., and Gao, Z. L. (2021). Carbon source applied in enrichment stage of mixed microbial cultures limits the substrate adaptability for PHA fermentation using the renewable carbon. *Appl. Biochem. Biotechnol.* 193, 3253–3270. doi:10.1007/s12010-021-03587-9

Zheng, Y., Chen, J. C., Ma, Y. M., and Chen, G. Q. (2020). Engineering biosynthesis of polyhydroxyalkanoates (PHA) for diversity and cost reduction. *Metab. Eng.* 58, 82–93. doi:10.1016/j.ymben.2019.07.004

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 08 February 2023
DOI 10.3389/fbioe.2023.1130405

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Zheng-Jun Li,
Beijing University of Chemical Technology, China
Yujia Jiang,
Nanjing Tech University, China

*CORRESPONDENCE
Chen-Guang Liu,
cg.liu@sjtu.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 23 December 2022
ACCEPTED 27 January 2023
PUBLISHED 08 February 2023

CITATION
Cao L-Y, Liu C-G, Yang S-H and Bai F-W (2023), Regulation of biofilm formation in *Zymomonas mobilis* to enhance stress tolerance by heterologous expression of *pfs* and *luxS*.
Front. Bioeng. Biotechnol. 11:1130405.
doi: 10.3389/fbioe.2023.1130405

Regulation of biofilm formation in *Zymomonas mobilis* to enhance stress tolerance by heterologous expression of *pfs* and *luxS*

Lian-Ying Cao[1], Chen-Guang Liu[1]*, Shi-Hui Yang[2] and Feng-Wu Bai[1]

[1]State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Science, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China, [2]State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan, China

Zymomonas mobilis is a potential alternative of *Saccharomyces cerevisiae* to produce cellulosic ethanol with strengths in cofactor balance, but its lower tolerance to inhibitors in the lignocellulosic hydrolysate restricts the application. Although biofilm can improve bacteria stress tolerance, regulating biofilm formation in *Z. mobilis* is still a challenge. In this work, we constructed a pathway by heterologous expressing *pfs* and *luxS* from *Escherichia coli* in *Z. mobilis* to produce AI-2 (autoinducer 2), a universal quorum-sensing signal molecule, to control cell morphology for enhancing stress tolerance. Unexpectedly, the results suggested that neither endogenous AI-2 nor exogenous AI-2 promoted biofilm formation, while heterologous expression of *pfs* can significantly raise biofilm. Therefore, we proposed that the main factor in assisting biofilm formation was the product accumulated due to heterologous expression of *pfs*, like methylated DNA. Consequently, ZM4::pfs produced more biofilm, which presented an enhanced tolerance to acetic acid. All these findings provide a novel strategy to improve the stress tolerance of *Z. mobilis* by enhancing biofilm formation for efficient production of lignocellulosic ethanol and other value-added chemical products.

KEYWORDS

***Zymomonas mobilis*, quorum sensing, *pfs*, *luxS*, biofilm, stress tolerance**

1 Introduction

Zymomonas mobilis is a promising strain in lignocellulosic ethanol production because of its rapid usage of glucose through the Entner–Doudoroff pathway and its superior ability to efficiently utilize pentose by expressing four genes encoding xylose assimilation and pentose phosphate pathway enzymes from *Escherichia coli* (Zhang et al., 1995; Ming et al., 2014). The toxic byproducts from lignocellulosic hydrolysate significantly impair the growth and viability of *Z. mobilis* (Franden et al., 2013), thus limits its application in industrial fermentation. Several traits, such as adaptive laboratory evolution, error-prone PCR-based whole genome shuffling, transposon-based random mutagenesis, omics mining, and molecular manipulation, have been successfully exploited to strengthen the tolerance of *Z. mobilis* to specific inhibitors (Yang et al., 2010a; Yang et al., 2010b; Jia et al., 2013; Shui et al., 2015; Wang et al., 2016; Huang et al., 2018). However, the requirement for tolerance to multiple inhibitors in lignocellulosic hydrolysate is still a challenge.

Biofilm is an ordered structure composed of microorganisms, extracellular polysaccharide substance (EPS), protein, and extracellular DNA, which could immobilize microorganisms within the bioreactor to improve fermentation efficiency (Todhanakasem et al., 2019; Xia et al.,

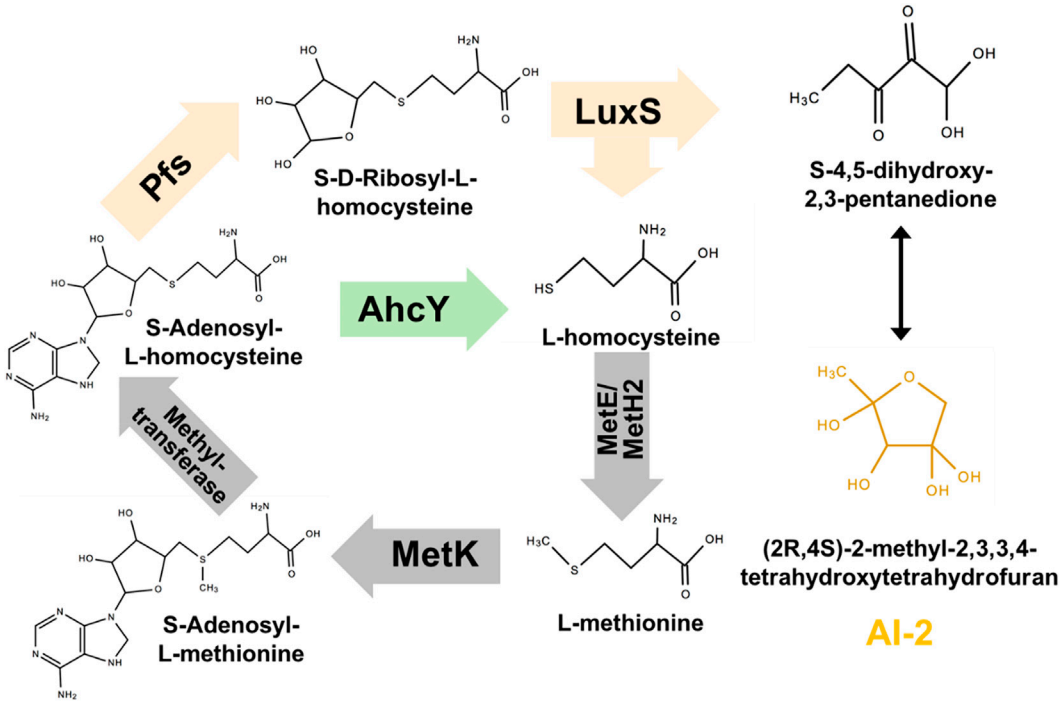


FIGURE 1
Constructing exogenous pathway for AI-2 synthesis in *Z. mobilis* ZM4. The orange genes were heterologously expressed in ZM4. The green gene should be deleted.

2019). Studies have shown that biofilm formation of *Z. mobilis* can significantly enhance the tolerance to various inhibitors compared with planktonic cells (Todhanakasem, 2016; Todhanakasem et al., 2019). Biotic or abiotic carriers such as corn silk, DEAE-cellulose, polystyrene, and polyvinyl chloride have been employed as platforms for *Z. mobilis* to generate biofilm with improved fermentation efficiency (Todhanakasem, 2016; Todhanakasem et al., 2019).

Quorum sensing (QS) is an intercellular communication triggered by a high dosage of specific organic molecules, usually causes acute morphological or physiological changes of bacteria such as bioluminescence, virulence factor secretion, biofilm formation, and biofilm dispersion (Whiteley et al., 2017). For decades, several sorts of QS signal molecules have been identified, including AI-2 (autoinducer-2), AHLs (acyl-homoserine lactone), DSFs (diffusible signal molecules), and AIP (autoinducing peptide) (Papenfort and Bassler, 2016). Among them, AI-2 is the most widespread QS signal molecule that exists in both Gram-positive and Gram-negative bacteria, mediates inter-species and intra-species communication (Pereira et al., 2013). The role of AI-2 to regulate biofilm formation has been recognized in various species (Yang, 2011), but its function in *Z. mobilis* has not been investigated.

The synthesis pathway of AI-2 is conserved among all microorganisms: as shown in Figure 1, AI-2 is generated by S-ribosylhomocysteinase (LuxS) from S-ribosylhomocysteine (SRH), which is produced by 5′-Methylthioadenosine/S-adenosylhomocysteine nucleosidase (Pfs) from S-Adenosyl-L-homocysteine (SAH) (Pereira et al., 2013). This two-step synthetic pathway of AI-2 is also a part of cysteine and methionine circulation, existing in *Gammaproteobacteria*, *Betaproteobacteria*, *Epsilonproteobacteria*, *Spirochaetes*, *Actinobacteria*, *Firmicutes*, and *Deinococcus-thermus* (Sun et al., 2004). Instead of the two-step way conducted by Pfs and LuxS, *Z. mobilis* uses one-step hydrolysis to break up S-Adenosyl-L-homocysteine (SAH) and to complete the methyl cycle, making the cell unable to generate AI-2 (Yang et al., 2009; Yang et al. 2018).

In this work, we constructed the AI-2 synthetic pathway in *Z. mobilis* by heterologous expressing *pfs* and *luxS* from *E. coli* K12 (MG1655). Although the reconstructed strain successfully synthesized AI-2, its function in promoting the biofilm formation was not observed. Surprisingly, expression of only *pfs* significantly enhanced biofilm formation, which enhanced stress tolerance of strain to acetic acid, a major inhibitor in the lignocellulosic hydrolysate.

2 Materials and methods

2.1 Strains, media, and cultivation

All strains used in this study are listed in Table 1. Wildtype ZM4 was purchased from American Type Culture Collection (ATCC) and cultivated in RMG (10 g/L yeast extract, 20 g/L glucose, and 2 g/L KH_2PO_4) medium. For seed culture preparation, one colony of *Z. mobilis* was picked up from a solid plate and inoculated into a 5 mL RMG. After 24 h cultivation, 1 mL OD_{600} = 1.5 cultures was transferred into 100 mL RMG and cultivated statically at 30°C. The medium for transformant strain was supplied with 20 mg/L tetracycline. *E. coli* DH5α and *E. coli* JM110 applied in this work were used for the construction and propagation of plasmids carrying genes with and without methylation, respectively, which were grown in 5 mL

TABLE 1 Plasmids used in this work.

Strains/Plasmids	Description
E. coli DH5α	*lacZΔM15, recA1*
E. coli JM110	*rpsL, dam-, dcm-*
Z. mobilis ZM4	ATCC31821
ZM4/pHW20a	ZM4 engineered with the vector pHW20a
ZM4/20a::pfs	ZM4 engineered with the overexpression of pfs from MG1655
ZM4/20a::luxS	ZM4 engineered with the overexpression of luxS from MG1655
ZM4/20a::pfs-luxS	ZM4 engineered with the overexpression of pfs and luxS from MG1655
E. coli K12 MG1655	ATCC 47076
Vibrio fisheri BB152	ATCC BAA-1117
Vibrio fisheri BB170	ATCC BAA-1119
pHW20a::P*gap*	pHW20a containing P*gap* from ZM4
pHW20a::pfs	pHW20a containing *pfs*of MG1655 driven by $Pgap^{Zm}$
pHW20a::luxS	pHW20a containing *luxS* of MG1655 driven by $Pgap^{Zm}$
pHW20a::pfs-luxS	pHW20a containing pfs and luxS of MG1655 driven by PgapZm
pEX18Tc	gene replacement vector with multi clone sites from pUC18
pEX18Tc?0182	pEX18Tc containing recombinant sequences flanking ZMO0182

TABLE 2 Primers used in this work.

Primer	Sequence (5′-3′)	Description
pfs-F	aataagttaggagaataaacATGAAAATCGGCATCATTGGTGCAA	Amplify *pfs* for construction of 20a::pfs
pfs-R	ctagaggatccccgggtaccTTAGCCATGTGCAAGTTTCTGCAC	
luxS-F	aataagttaggagaataaacATGCCGTTGTTAGATAGCTTCACAG	Amplify luxS for construction of 20a::luxS
luxS-R	ctagaggatccccgggtaccCTAGATGTGCAGTTCCTGCAACTTC	
pfs-F	aataagttaggagaataaacATGAAAATCGGCATCATTGGTGCAA	Amplify pfs and luxS for construction of 20a::pfs-luxS
pfs (-luxS)-R	TTAGCCATGTGCAAGTTTCTGCAC	
(pfs-)luxS-F	AGAAACTTGCACATGGCTAAATGCCGTTGTTAGATAGCT TCACAG	
luxS-R	ctagaggatccccgggtaccCTAGATGTGCAGTTCCTGCAACTTC	
0182-HI-F	ttgcatgcctgcaggtcgactctagaCGAAGGCAGGCTGCCCCTGC	Amplify homologous fragments flank ZMO0182 for construction of pEX18Tc? ZMO0182
0182-H1-R	AGCAAGGCTGATGTCACGGA	
0182-H2-F	TCCGTGACATCAGCCTTGCTGGATCATTATCGTTATTGATT	
0182-H2-R	aattcgagctcggtacccggggatccTCGCTACCCGCGCTTATGTC	

Luria-Bertani (5 g/L yeast extract, 10 g/L tryptone, 10 g/L sodium chloride) of 15 mL tubes at 37°C, 200 rpm. *V. fisheri* BB152 and *V. fisheri* BB170, donated by Prof. Xiangan Han (Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences) and Prof. Bonnie Bassler (Princeton University), were recovered in MB (Marine Broth 2216) and cultivated in AB (Autoinducer Bioassay) medium (2.0 g/L casamino acids, vitamin-free, 12.3 g/L magnesium sulfate heptahydrate, 17.5 g/L sodium chloride, 174 mg/L L-arginine, 1% (v/v) glycerol and 10 mM potassium phosphate) for AI-2 detection.

2.2 Construction of AI-2 synthetic pathway

Pfs can catalyze SAH to SRH, and LuxS splits SRH to L-homocysteine and S-4,5-dihydroxy-2,3-pentanedione (auto-

cyclizing into AI-2) (Figure 1). With the transformation of *pfs* and *luxS*, *Z. mobilis* can recover methionine and produce AI-2 simultaneously. SAH hydrolase AhcY (ZMO0182) in *Z. mobilis* hydrolyzes SAH to L-homocysteine directly without AI-2 production. Knock out of AhcY was conducted to redistribute more carbon flux into the constructed pfs-luxS pathway.

2.3 Construction and identification of transformants

Plasmids and primers used in this study are given in Tables 1, 2. For heterologous gene expression, the gap^{Zm} promoter was amplified by PCR and ligated with shuttle vector pHW20a in advance. Target genes were then amplified by PCR and infused with linearized vectors with $Pgap^{Zm}$ by seamless cloning (Seamless Cloning master mix, Sangon Biotech, Shanghai, China), followed by propagating into *E. coli* DH5α. Due to the Restriction-Modification systems in *Z. mobilis*, the confirmed plasmids by PCR and Sanger sequencing were further transformed into *E. coli* JM110 for demethylation. The demethylated plasmids were then extracted from *E. coli* JM110 and electro-transformed into ZM4 by the Gene Pulser (Gene Pulser Xcell™, Bio-Rad). Electro-transformation was done with 1 mm gap cuvettes operated at 1.8 kV. Colonies were selected by tetracycline and confirmed by PCR.

Gene deletion was performed by homologous recombination with the suicide vector pEX18Tc bearing tetracycline selection marker and sucrose counter-selection marker. Briefly, 500–1,000 bp fragments flanking ZMO0182 were amplified, and fused by seamless cloning as described above, which was propagated in *E. coli* DH5α, and confirmed by PCR and sequencing. The correct plasmid was transformed into *E. coli* JM110 to demethylation. Demethylated plasmid was electro-transformed into ZM4 as described above. The positive selection was screened using RMG supplemented with tetracycline after the first crossover recombination, followed by the second crossover recombination. A mutant with the gene deleted should be selected through the counter-selection using the rich medium supplemented with sucrose (Xia et al., 2019).

2.4 AI-2 detection

The process to detect AI-2 was modified based on previous work (Taga and Xavier, 2011). Briefly, the seed culture of *Z. mobilis* was inoculated into RMG with starting OD_{600} at 0.015 and cultivated at 30°C, 150 rpm for 18 h to enter the post-exponential phase when the most abundant AI-2 could accumulate. Triplicate samples, each with 4 mL, were centrifuged by Xiangyi H1650-W at 5,000 rpm for 5 min to collect the supernatant and passed it through a 0.22 μm sterile syringe filter. *Vibrio harveyi* BB170 was inoculated in a 15 mL flask with 5 mL AB medium overnight at 28°C, 150 rpm, and then diluted with AB medium at 1:5,000 as the seed culture. 0.2 mL supernatant and 1.8 mL seed culture were mixed and cultivated at 28°C, 150 rpm for AI-2 detection. The mixture's bioluminescence intensity was measured after 6 h by Multimode Plate Reader (PE & ENSPIRE 2300, Perkin-Elmer, United States) to indicate AI-2 concentration. Fresh AB medium was used to replace the sample as the negative control, and synthetic AI-2 (D060111, Omm Scientific) or supernatant of *V. Harveyi* BB152 were applied as the positive controls. Considering the inhibition of tetracycline on growth and bioluminescence of BB170 during AI-2 detection and the high stability of the pHW20a plasmid (Dong et al., 2011), all transformants were cultivated without antibiotic for supernatant collection.

2.5 Motility ration

Bacterial motility was displayed as the motility ratio, which reflected a height ratio of the turbid part in the whole medium in a tube. Briefly, the seed culture of *Z. mobilis* was inoculated into 5 mL RMG of 15 mL tubes and cultivated at 30°C statically. Cells with weak motility would settle down due to the gravity, while cells with high motility will swim through the entire tube to remain turbid. After 24 h, the height of the whole media and the transparent portion were measured. The motility ratio was calculated as (1- the height of the transparent part/the height of the entire medium).

2.6 Quantitative analysis of biofilm

For biofilm establishment, plastic (Polymethyl Methacrylate, PM) flake at 1 cm × 1 cm was chosen as the abiotic platform. 24-wells plate with PM flakes pre-loaded was inoculated with seed culture and incubated at 30°C statically for cell attachment. Biofilm development was visualized and quantified by crystal violet staining. After cultivation in wells for 3 days, plastic flakes with biofilm formed were picked out and soaked in 1% crystal violet for 20 min, followed by rinsing with deionized water twice. The remaining biofilm with crystal violet staining was dissolved in 1 mL 95% ethanol and measured at OD_{595} by spectrophotometer (Multiskan GO 1510, ThermoFisher, Finland). Biomass in 24-wells plate was collected by resuspension and then measured at OD_{600} by spectrophotometer.

2.7 Toxicity study

Acetic acid toxicity for biofilms was studied by LIVE/DEAD™ BacLight™ Bacterial Viability Kit (L7012, Invitrogen, Thermo Fisher Scientific, United States). Based on previous work (Li et al., 2006), cells grown on plastic flakes in wells for 3 days to form mature biofilm and then incubated at 21 g/L acetic acid for 0.5 h. Then flakes with biofilm were picked out and rinsed in sterile H_2O twice for being stained with the LIVE/DEAD Bac Light™ Bacterial Viability and Counting Kit, followed by examination with Super-resolution Multiphoton Confocal Microscope (TCS SP8 STED 3X, Leica, Germany).

2.8 Test of stress tolerance

10 mL OD_{600} = 1.5 seed culture was inoculated into 90 mL RMG. The culture was cultivated at 30°C, 150 rpm for 12 h to measure OD_{600}. The strains' tolerance to 7.5% ethanol, 2.1 g/L acetic acid (density: 1.05 g/mL), or 1 g/L vanillin were evaluated by comparing with OD_{600} of cells cultivated without inhibitor. Transformants were supplemented with 20 mg/L tetracycline.

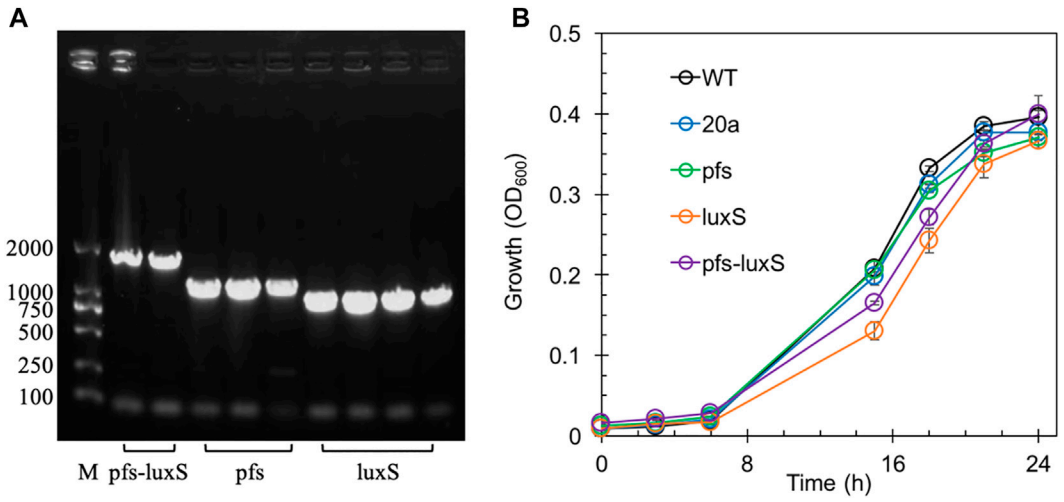


FIGURE 2
The confirmation of genes manipulation **(A)** and growth profiles of transformants and wildtype *Z. mobilis* **(B)**. M stands for DNA marker DL 2000.

3 Results and discussion

3.1 Growth of strains with heterologous gene expression

Transformants with heterologous expression of *pfs* and *luxS* were selected by tetracycline screen and PCR confirmation. Figure 2A showed the correct length of bands obtained by colony PCR of transformants, hinting that *pfs* or/and *luxS* were successfully transformed into ZM4. Knock out of AhcY was also conducted to enhance the Pfs-LuxS pathway. However, no transformant with AhcY deletion was obtained, which might because S-adenosyl-L-homocysteine hydrolase AhcY is essential for *Z. mobilis* to complete the cysteine and methionine circulation and prevent the accumulation of toxicant SAH in wildtype. To elucidate the effects of Pfs and LuxS on growth, the biomass of strains was evaluated before the investigation of other physiological performances. As visualized in Figure 2B, ZM4:luxS and ZM4:pfs-luxS had a slower growth rate in the log phase (6–20 h) when compared with WT and ZM4:20a, which suggested that the expression of *luxS* might cause a metabolic burden on bacteria by plasmid maintenance and heterologous gene expression (Pasini et al., 2016). However, heterologous expression of *pfs* leads to almost no metabolic burden in terms of growth rate, which may occur due to the clearance of toxic SAH by Pfs (Parveen and Cornell, 2011). All strains approached similar OD_{600} at stationary phase 24 h, indicating that the side effect of heterologous expression of *pfs* and/or *luxS* was negligible at the end of growth (less than 5%), eliminates the requirement for considering the metabolic burden in the following experiments.

3.2 AI-2 produced by expression of *pfs* and *luxS*

In principle, ZM4 with *pfs* and *luxS* heterologous expression could produce AI-2 even though with the failure deletion of ZMO0182. The bioluminescent density of *V. fisheri* BB170 was used to semi-quantify AI-2 when taking AI-2 and *V. fisheri* BB152 supernatant as positive controls and AB medium as the negative control. As shown in Figure 3, the AB medium without AI-2 allowed BB170 to present the basal bioluminescence. The density of bioluminescence positively correlated with the concentration of AI-2. Based on the relation, the AI-2 in BB152 supernatant should go below 4.47 μM. Interestingly, the supernatant of *Z. mobilis* strains except for ZM4:pfs-luxS pulled down the bioluminescence of BB170, making it lower than the negative control, which was reasonable as the metabolites of *Z. mobilis*, such as major product ethanol, possibly played an inhibitory role on

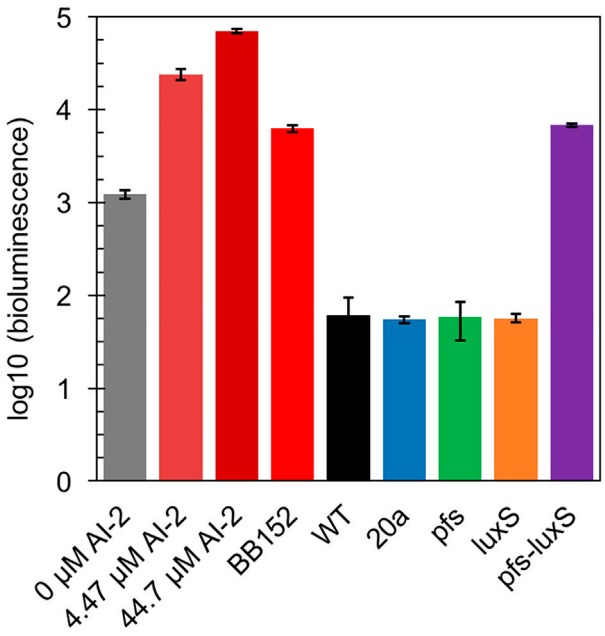


FIGURE 3
Bioluminescence assay by *V. fisheri* BB170 for AI-2 semi-quantification in the supernatant of cultures.

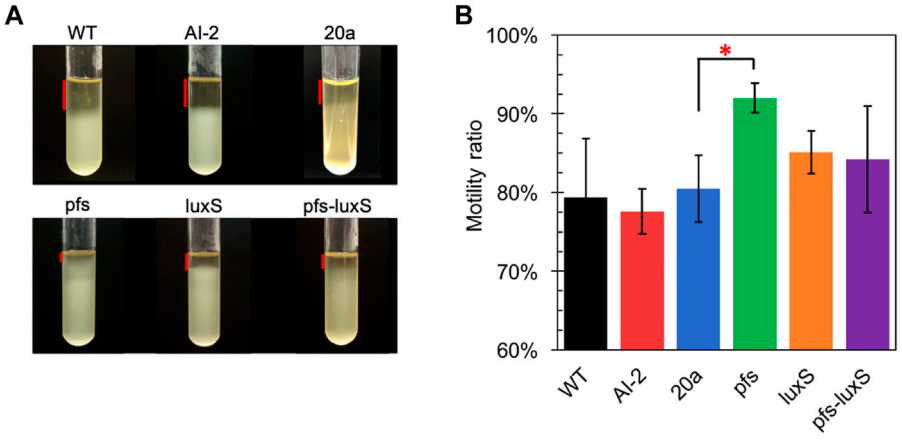


FIGURE 4
Motility test of *Z. mobilis* strains by settlement method **(A)** The cell sedimentation in the tubes **(B)** Motility ratio = 1-the height of the transparent part/the height of the entire medium. WT: wildtype. All transformants are constructed from ZM4. Three replicates. * represents p-value < 0.1.

bioluminescent. Even though, the supernatant of ZM4:pfs-luxS lead to a higher bioluminescence intensity than the negative control and other *Z. mobilis* strains, which demonstrated that AI-2 was successfully produced and the functions of Pfs and LuxS can be confirmed in *Z. mobilis*.

3.3 Motility strengthened by *pfs* in an AI-2 independent way

Many studies regarding AI-2-mediated quorum sensing have confirmed its role in motility, which includes swimming motility and swarming motility (Teren et al., 2018; Wang et al., 2018). The test of cell motility using static cultivation in tubes reveals the relationship between quorum sensing and the motility of ZM4. The sedimentation of cells during static cultivation can be interfered with by the vigorous motility of *Z. mobilis*, resulting in an unevenly distributed transparent medium in the upper part and turbid culture in the lower part (Figure 4).

Figure 4 showed that wildtype owned a low motility ratio, which was not changed, even adding 4.47 μM AI-2, which was straight evidence to confirm the invalidation of AI-2 on the mobility of *Z. mobilis*. Comparing with wildtype, ZM4:20a showed no statistically different motility, but the heterologous expression of *pfs* strengthened cell motility. Interestingly, neither ZM4:luxS nor ZM4:pfs-luxS performed significantly enhanced mobility, even though the AI-2 produced by ZM4: pfs-luxS, reflected that the mobility of *Z. mobilis* depends on other specific reactions catalyzed by Pfs, instead of QS signal molecules AI-2. Pfs promotes SAM-dependent transmethylation by reducing SAH, simultaneously yield methylated DNA, RNA, and protein (Parveen and Cornell, 2011). Considering DNA methylation could regulate expression of motility-related genes (Vandenbussche et al., 2020), it is reasonable that strain with heterologous expression of *pfs* exhibited the strongest motility. ZM4:pfs-luxS also expressed Pfs and showed slightly enhanced motility but with insignificance, which might arise due to the relatively low expression of *pfs* as it was co-expressed with *luxS*.

3.4 Biofilm enhanced by *pfs* in an AI-2 independent way

Altered motility usually affects biofilm formation (Tolker-Nielsen, 2015; Jani et al., 2017). AI-2 has been verified to promote cell aggregation and biofilm formation in various bacteria (Wang et al., 2014; Laganenka et al., 2016; Mizan et al., 2016), but its role in *Z. mobilis* remains elusive. *Z. mobilis* exhibited greater attachment and improved biofilm formation on the hydrophobic surface than on hydrophilic glass in this study. The biofilm of wildtype and AI-2 producing strain ZM4: pfs-luxS were quantified to investigate the role of AI-2 on biofilm formation. As shown in Figure 5A, bacteria kept growing in the first 3 days and then entered the stationary phase. Thus, the third day with the highest OD_{600} was chosen for biofilm quantification.

Figure 5 shows a comparable quantity of biofilm formed by strains with or without AI-2, which reflects that the exogenous addition of AI-2 was incapable of contributing to biofilm formation in *Z. mobilis*. Moreover, the biofilm formation of ZM4:pfs-luxS and ZM4:20a were at the same level, presenting that the endogenous production of AI-2 did not correlate with the biofilm formation of *Z. mobilis*. ZM4:pfs produced the most abundant biofilm, which suggested that the product of Pfs might play a critical role in biofilm formation in *Z. mobilis*. The large error bar in Figure 5C indicates that the biofilm of *Z. mobilis* was not as tight as other bacteria biofilms such as *Pseudomonas aeruginosa* or *Bacillus subtilis*, which could be easily removed through rinsing (Figure 5B) (Chua et al., 2014; Duanisassaf et al., 2016). Researches on *E. coli* revealed that the cell's high motility contributes to forming surface-adherent structures (Jani et al., 2017). Combining this issue with the previous motility evaluation, since adherence of cells to the plastic surface is induced by physical contact, ZM4:pfs with higher motility hold a greater probability of contact to the plastic surface in more biofilm formation. Plus, methylated DNA has been reported to play an essential role in efficient biofilm formation (Aya Castañeda et al., 2015) which could explain the increased volume of biofilm formed by ZM4:pfs.

Typically, biofilm formation consists of five steps: attachment, aggregation, maturation, full development, and dispersion (Verderosa et al., 2019). The first step when cells migrate and adhere to an abiotic

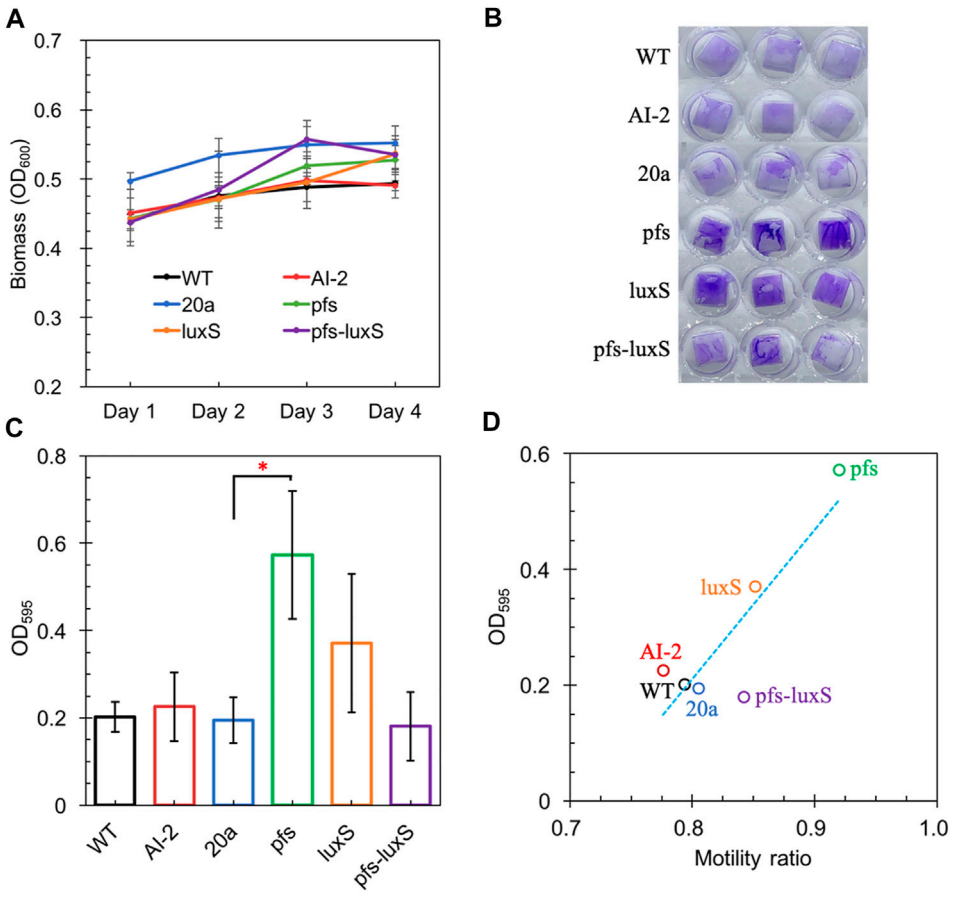


FIGURE 5
The Growth of strains on a 24-wells plate **(A)**; Crystal violet staining of WT, WT (with 4.47 μM AI-2 addition), ZM4:20a, ZM4:pfs, ZM4:luxS, and ZM4:pfs-luxS **(B)**; Quantification of biofilm **(C)**; Correlations between biofilm and bacterial motility **(D)**. Three replicates. * represents p-value < 0.1.

surface is severely affected by motility, which is in line with the strong positive correlation between biofilm formation and motility observed in our work (Figure 5D), hinted this step might be the most influential one during biofilm formation by *Z. mobilis*.

3.5 Stress tolerance of *Zymomonas mobilis* strengthened by biofilm

Biofilm formation has been proven to enhance stress tolerance in several *Z. mobilis* strains (Li et al., 2006; Todhanakasem et al., 2018). Studies have confirmed that biofilm can help *Z. mobilis* be more tolerant to various inhibitors in lignocellulosic hydrolysate (Li et al., 2006; Todhanakasem et al., 2014). We measured the biofilm strains' viability to test their stress tolerance to acetic acid, one of the significant lignocellulosic hydrolysate inhibitors. Figure 6 showed that the viability of ZM4:pfs was much higher than WT, which indicated increased biofilm formed by ZM4:pfs was helpful to resist the stress from inhibitor outside. The results verified the role of accumulated biofilm in improving tolerance of bacteria to inhibitors.

As ZM4:pfs had greater tolerance to acetic acid, but it remains uncertain whether this strengthens derived from the enzyme Pfs or the protection by biofilm. Fermentation with inhibitors was conducted to

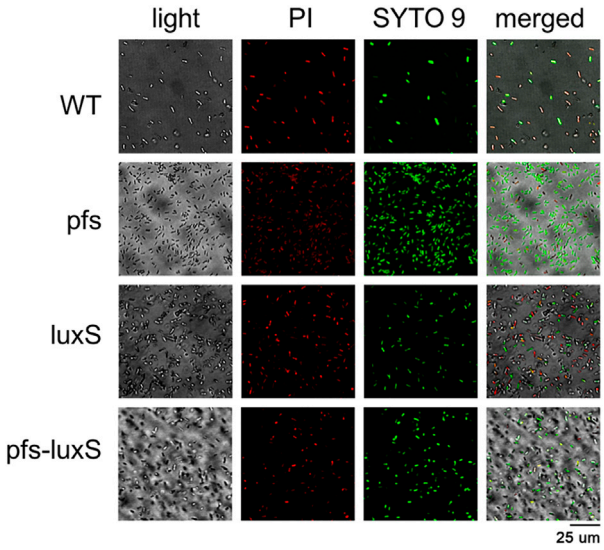


FIGURE 6
Images of WT, ZM4:pfs, ZM4:luxS, ZM4:pfs-luxS after 2.1 g/L acetic acid shock. SYTO 9 stain and propidium iodide were used to stain cells. Live cells fluoresced green and dead cells fluoresced red.

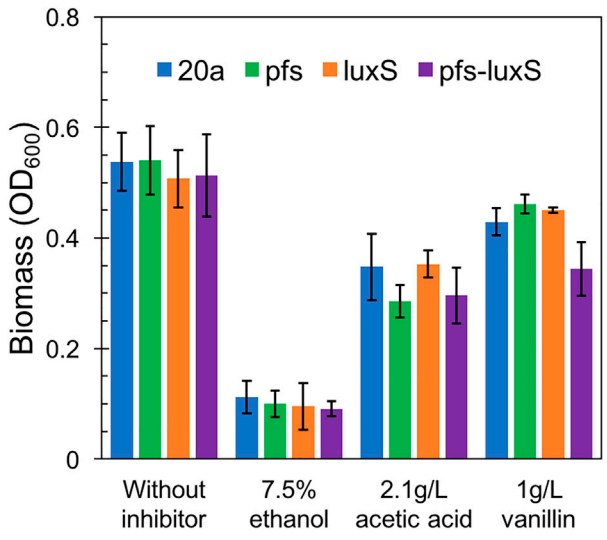


FIGURE 7
Growth of planktonic cells in flasks under ethanol, acetic acid, and vanillin addition, OD_{600} was measured at 12 h. The statistic *p*-value >0.05 among each group.

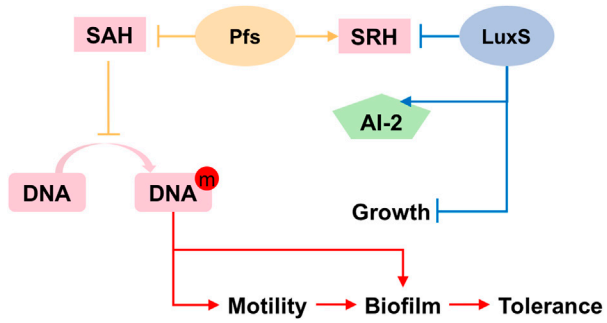


FIGURE 8
Proposed mechanism of Pfs and LuxS on AI-2 production, motility, biofilm formation, and stress tolerance.

address this question. As Figure 7 showed, planktonic cells reached similar OD at 12 h without inhibitor in the rotated flask, whose hydrophobic glass surface would stop the biofilm formation. ZM4:pfs performed a similar growth with ZM4:pHW20a under the stressful conditions with ethanol, acetic acid, or vanillin supplementation. These results confirmed the improvement of stress tolerance in ZM4:pfs was due to an increase in the quantity of biofilm induced by Pfs instead of the product of protein Pfs.

3.6 The mechanism of pfs and LuxS on biofilm formation and stress tolerance

Based on the results illustrated above, the mechanism underlying the heterologous expression of *pfs* and *luxS* is proposed. As Figure 8 shows, the co-expression of *pfs* and *luxS* enables ZM4 to produce AI-2. However, AI-2 does not affect the growth, motility, biofilm formation, or stress tolerance of ZM4. Unexpectedly, expression of *pfs* alone enhances cell motility and biofilm formation, thus strengthens strains' stress tolerance, which might due to the accumulated methylated DNA because of the inhibition by low concentration of SAH that was utilized by Pfs (Aya Castañeda et al., 2015). Whereas, co-expression of *pfs* and *luxS* do not show similar results as the cysteine-methionine circulation is finished so that SAH cannot be maintained at low concentration and no extra methylated DNA could accumulate. Besides, Pfs can hydrolyze SAH into SRH, which might contribute to biofilm formation (Challan Belval et al., 2006). Finally, increased biofilm formation protects cells from inhibitors added exogenously, thus strengthened bacterial stress tolerance.

4 Conclusion

We constructed an AI-2 producing strain by heterologous expression of *pfs* and *luxS* with a negligible metabolic burden. However, neither cell motility nor biofilm formation was observed to be improved by either endogenously generated or exogenously added AI-2. Surprisingly, expressing *pfs* alone enhanced cell motility and biofilm formation, contributing to stress tolerance strengthen. This work verifies the importance of biofilm on environmental stresses and provides a new method to improve biofilm formation in *Z. mobilis*.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

L-YC: Investigation, Writing-original draft preparation; C-GL: Conceptualization, Supervision, Writing-reviewing and editing; S-HY: Writing-reviewing, Resources; F-WB: Methodology, Writing-reviewing, Supervision. All authors contributed to the article and approved the submitted version.

Funding

We appreciate financial support from grants sponsored by National Key Research and Development Program of China from Ministry of Science and Technology of the People's Republic of China (2021YFC2101300), National Natural Science Foundation of China (NSFC) (31970026 and 21978167). Open Project Funding of the State Key Laboratory of Biocatalysis and Enzyme Engineering (SKLBEE2018016).

Acknowledgments

We also thank the Core Facility and Service Center (CFSC) in School of Life Sciences and Biotechnology, SJTU for metabolite analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aya Castañeda, M. D. R., Sarnacki, S. H., Noto Llana, M., López Guerra, A. G., Giacomodonato, M. N., and Cerquetti, M. C. (2015). Dam methylation is required for efficient biofilm production in *Salmonella enterica* serovar *Enteritidis*. *Int. J. Food Microbiol.* 193, 15–22. doi:10.1016/j.ijfoodmicro.2014.10.003

Challan Belval, S., Gal, L., Margiewes, S., Garmyn, D., Piveteau, P., and Guzzo, J. (2006). Assessment of the roles of LuxS, S-ribosyl homocysteine, and autoinducer 2 in cell attachment during biofilm formation by *Listeria monocytogenes* EGD-e. *Appl. Environ. Microbiol.* 72 (4), 2644–2650. doi:10.1128/AEM.72.4.2644-2650.2006

Chua, S. L., Liu, Y., Yam, J. K., Chen, Y., Vejborg, R. M., Tan, B. G., et al. (2014). Dispersed cells represent a distinct stage in the transition from bacterial biofilm to planktonic lifestyles. *Nat. Commun.* 5, 4462. doi:10.1038/ncomms5462

Dong, H. W., Bao, J., Ryu, D. D., and Zhong, J. J. (2011). Design and construction of improved new vectors for *Zymomonas mobilis* recombinants. *Biotechnol. Bioeng.* 108, 1616–1627. doi:10.1002/bit.23106

Duanisassaf, D., Steinberg, D., Chai, Y., and Shemesh, M. (2016). The LuxS based quorum sensing governs lactose induced biofilm formation by *Bacillus subtilis*. *Front. Microbiol.* 6, 1517. doi:10.3389/fmicb.2015.01517

Franden, M. A., Pilath, H. M., Mohagheghi, A., Pienkos, P. T., and Zhang, M. (2013). Inhibition of growth of *Zymomonas mobilis* by model compounds found in lignocellulosic hydrolysates. *Biotechnol. Biofuels* 6 (1), 99. doi:10.1186/1754-6834-6-99

Huang, S., Xue, T., Wang, Z., Ma, Y., He, X., Hong, J., et al. (2018). Furfural-tolerant *Zymomonas mobilis* derived from error-prone PCR-based whole genome shuffling and their tolerant mechanism. *Appl. Microbiol. Biotechnol.* 102 (7), 3337–3347. doi:10.1007/s00253-018-8817-8

Jani, S., Seely, A. L., Peabody, V. G. L., Jayaraman, A., and Manson, M. D. (2017). Chemotaxis to self-generated AI-2 promotes biofilm formation in *Escherichia coli*. *Microbiology* 163 (12), 1778–1790. doi:10.1099/mic.0.000567

Jia, X., Wei, N., Wang, T., and Wang, H. (2013). Use of an EZ-Tn5-based random mutagenesis system to create a *Zymomonas mobilis* with significant tolerance to heat stress and malnutrition. *J. Ind. Microbiol. Biotechnol.* 40 (8), 811–822. doi:10.1007/s10295-013-1287-1

Laganenka, L., Colin, R., and Sourjik, V. (2016). Chemotaxis towards autoinducer 2 mediates autoaggregation in *Escherichia coli*. *Nat. Commun.* 7, 12984. doi:10.1038/ncomms12984

Li, X. Z., Webb, J. S., Kjelleberg, S., and Rosche, B. (2006). Enhanced benzaldehyde tolerance in *Zymomonas mobilis* biofilms and the potential of biofilm applications in fine-chemical production. *Appl. Environ. Microbiol.* 72 (2), 1639–1644. doi:10.1128/AEM.72.2.1639-1644.2006

Ming, X. H., Bo, W., Han, Q., Zhi, Y. R., Fu, R. T., Wang, J., et al. (2014). *Zymomonas mobilis*: A novel platform for future biorefineries. *Biotechnol. Biofuels* 7, 101. doi:10.1186/1754-6834-7-101

Mizan, M. F. R., Jahid, I. K., Kim, M., Lee, K.-H., Kim, T. J., and Ha, S.-D. (2016). Variability in biofilm formation correlates with hydrophobicity and quorum sensing among Vibrio parahaemolyticus isolates from food contact surfaces and the distribution of the genes involved in biofilm formation. *Biofouling* 32 (4), 497–509. doi:10.1080/08927014.2016.1149571

Papenfort, K., and Bassler, B. L. (2016). Quorum sensing signal-response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* 14 (9), 576–588. doi:10.1038/nrmicro.2016.89

Parveen, N., and Cornell, K. A. (2011). Methylthioadenosine/S-adenosylhomocysteine nucleosidase, a critical enzyme for bacterial metabolism. *Mol. Microbiol.* 79 (1), 7–20. doi:10.1111/j.1365-2958.2010.07455.x

Pasini, M., Fernández-Castané, A., Jaramillo, A., De Mas, C., Caminal, G., and Ferrer, P. (2016). Using promoter libraries to reduce metabolic burden due to plasmid-encoded proteins in recombinant *Escherichia coli*. *Escherichia Coli. N. Biotechnol.* 33 (1), 78–90. doi:10.1016/j.nbt.2015.08.003

Pereira, C. S., Thompson, J. A., and Xavier, K. B. (2013). AI-2-mediated signalling in bacteria. *FEMS Microbiol. Rev.* 37 (2), 156–181. doi:10.1111/j.1574-6976.2012.00345.x

Shui, Z. X., Qin, H., Wu, B., Ruan, Z. Y., Wang, L. S., Tan, F. R., et al. (2015). Adaptive laboratory evolution of ethanologenic *Zymomonas mobilis* strain tolerant to furfural and acetic acid inhibitors. *Appl. Microbiol. Biotechnol.* 99 (13), 5739–5748. doi:10.1007/s00253-015-6616-z

Sun, J., Daniel, R., Wagnerdöbler, I., and Zeng, A. P. (2004). Is autoinducer-2 a universal signal for interspecies communication: A comparative genomic and phylogenetic analysis of the synthesis and signal transduction pathways. *BMC Evol. Biol.* 4, 36. doi:10.1186/1471-2148-4-36

Taga, M. E., and Xavier, K. B. (2011). Methods for analysis of bacterial autoinducer-2 production. *Curr. Protoc. Microbiol.* 23. Chapter 1: Unit 1C.1. doi:10.1002/9780471729259.mc01c01s23

Teren, M., Turonova Michova, H., Vondrakova, L., and Demnerova, K. (2018). Molecules autoinducer 2 and cjA and their impact on gene expression in *Campylobacter jejuni*. *J. Mol. Microbiol. Biotechnol.* 28 (5), 207–215. doi:10.1159/000495411

Todhanakasem, T., Salangsing, O. L., Koomphongse, P., Kaewket, S., Kanokratana, P., and Champreda, V. (2019). *Zymomonas mobilis* biofilm reactor for ethanol production using rice straw hydrolysate under continuous and repeated batch processes. *Front. Microbiol.* 10, 1777. doi:10.3389/fmicb.2019.01777

Todhanakasem, T., Sangsutthiseree, A., Areerat, K., Young, G. M., and Thanonkeo, P. (2014). Biofilm production by *Zymomonas mobilis* enhances ethanol production and tolerance to toxic inhibitors from rice bran hydrolysate. *N. Biotechnol.* 31 (5), 451–459. doi:10.1016/j.nbt.2014.06.002

Todhanakasem, T., Yodsanga, S., Sowatad, A., Kanokratana, P., Thanonkeo, P., and Champreda, V. (2018). Inhibition analysis of inhibitors derived from lignocellulose pretreatment on the metabolic activity of *Zymomonas mobilis* biofilm and planktonic cells and the proteomic responses. *Biotechnol. Bioeng.* 115 (1), 70–81. doi:10.1002/bit.26449

Todhanakasem, T. (2016). *Zymomonas mobilis* biofilm enhances bioethanol production from lignocellulosic hydrolysate. *N. Biotechnol.* 33, S90–S91. doi:10.1016/j.nbt.2016.06.1035

Tolker-Nielsen, T. (2015). Biofilm development. *Microbiol. Spectr.* 3 (2), 0001–2014. MB-0001. doi:10.1128/microbiolspec.MB-0001-2014

Vandenbussche, I., Sass, A., Pinto-Carbó, M., Mannweiler, O., Eberl, L., and Coenye, T. (2020). DNA methylation epigenetically regulates gene expression in *Burkholderia cenocepacia* and controls biofilm formation, cell aggregation, and motility. *mSphere* 5 (4), 004555–e520. doi:10.1128/mSphere.00455-20

Verderosa, A. D., Totsika, M., and Fairfull-Smith, K. E. (2019). Bacterial biofilm eradication agents: A current review. *Front. Chem.* 7, 824. doi:10.3389/fchem.2019.00824

Wang, J. L., Wu, B., Qin, H., You, Y., Liu, S., Shui, Z. X., et al. (2016). Engineered *Zymomonas mobilis* for salt tolerance using EZ-Tn5-based transposon insertion mutagenesis system. *Microb. Cell Fact.* 15 (1), 101. doi:10.1186/s12934-016-0503-x

Wang, Y., Wang, Y., Sun, L., Grenier, D., and Yi, L. (2018). The LuxS/AI-2 system of *Streptococcus suis*. *Appl. Microbiol. Biotechnol.* 102 (17), 7231–7238. doi:10.1007/s00253-018-9170-7

Wang, Y., Yi, L., Zhang, Z., Fan, H., Cheng, X., and Lu, C. (2014). Biofilm formation, host-cell adherence, and virulence genes regulation of *Streptococcus suis* in response to autoinducer-2 signaling. *Curr. Microbiol.* 68 (5), 575–580. doi:10.1007/s00284-013-0509-0

Whiteley, M., Diggle, S. P., and Greenberg, E. P. (2017). Progress in and promise of bacterial quorum sensing research. *Nature* 551 (7680), 313–320. doi:10.1038/nature24624

Xia, J., Yang, Y., Liu, C. G., Yang, S. H., and Bai, F. W. (2019). Engineering *Zymomonas mobilis* for robust cellulosic ethanol production. *Trends Biotechnol.* 37 (9), 960–972. doi:10.1016/j.tibtech.2019.02.002

Yang, J. (2011). *Enhanced bioethanol production by* Zymomonas mobilis *in response to the quorum sensing molecules AI-2*. Durham, England: Durham University.

Yang, S. H., Land, M. L., Klingeman, D. M., Pelletier, D. A., Lu, T. Y., Martin, S. L., et al. (2010a). Paradigm for industrial strain improvement identifies sodium acetate tolerance loci in *Zymomonas mobilis* and *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 107 (23), 10395–10400. doi:10.1073/pnas.0914506107

Yang, S. H., Pappas, K. M., Hauser, L. J., Land, M. L., Chen, G. L., Hurst, G. B., et al. (2009). Improved genome annotation for *Zymomonas mobilis*. *Nat. Biotechnol.* 27 (10), 893–894. doi:10.1038/nbt1009-893

Yang, S. H., Pelletier, D. A., Lu, T. Y. S., and Brown, S. D. (2010b). The *Zymomonas mobilis* regulator hfq contributes to tolerance against multiple lignocellulosic pretreatment inhibitors. *BMC Microbiol.* 10, 135. doi:10.1186/1471-2180-10-135

Yang, S. H., Vera, J. M., Grass, J., Savvakis, G., Moskvin, O. V., Yang, Y. F., et al. (2018). Complete genome sequence and the expression pattern of plasmids of the model ethanologen *Zymomonas mobilis* ZM4 and its xylose-utilizing derivatives 8b and 2032. *Biotechnol. Biofuels.* 11, 125. doi:10.1186/s13068-018-1116-x

Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., and Picataggio, S. (1995). Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. *Science* 267 (5195), 240–243. doi:10.1126/science.267.5195.240

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 20 February 2023
DOI 10.3389/fbioe.2023.1141272

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Hui-Min Qin,
Tianjin University of Science and Technology, China
Shuli Liang,
South China University of Technology, China

*CORRESPONDENCE
Mingdong Yao,
mingdong.yao@tju.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 10 January 2023
ACCEPTED 06 February 2023
PUBLISHED 20 February 2023

CITATION
Zhang C, Chen W, Dong T, Wang Y, Yao M, Xiao W and Li B (2023), Elimination of enzymes catalysis compartmentalization enhancing taxadiene production in *Saccharomyces cerevisiae*. *Front. Bioeng. Biotechnol.* 11:1141272. doi: 10.3389/fbioe.2023.1141272

Elimination of enzymes catalysis compartmentalization enhancing taxadiene production in *Saccharomyces cerevisiae*

Chenglong Zhang[1], Wang Chen[1], Tianyu Dong[1], Ying Wang[1], Mingdong Yao[1]*, Wenhai Xiao[1,2] and Bingzhi Li[1]

[1]Frontier Science Center for Synthetic Biology, Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [2]Georgia Tech Shenzhen Institute, Tianjin University, Shenzhen, China

Taxadiene is an important precursor in taxol biosynthesis pathway, but its biosynthesis in eukaryotic cell factories is limited, which seriously hinders the biosynthesis of taxol. In this study, it is found that there was the catalysis compartmentalization between two key exogenous enzymes of geranylgeranyl pyrophosphate synthase and taxadiene synthase (TS) for taxadiene synthesis progress, due to their different subcellular localization. Firstly, the enzyme-catalysis compartmentalization was overcome by means of the intracellular relocation strategies of taxadiene synthase, including N-terminal truncation of taxadiene synthase and enzyme fusion of GGPPS-TS. With the help of two strategies for enzyme relocation, the taxadiene yield was increased by 21% and 54% respectively, among them the GGPPS-TS fusion enzyme is more effective. Further, the expression of GGPPS-TS fusion enzyme was improved *via* the multi-copy plasmid, resulting that the taxadiene titer was increased by 38% to 21.8 mg/L at shake-flask level. Finally, the maximum taxadiene titer of 184.2 mg/L was achieved by optimization of the fed-batch fermentation conditions in 3 L bioreactor, which is the highest reported titer of taxadiene biosynthesis accomplished in eukaryotic microbes. This study provides a successful example for improving biosynthesis of complex natural products by solving the critical problem of multistep enzymes catalysis compartmentalization.

KEYWORDS

taxadiene, catalysis compartmentalization, enzyme fusion, relocalization, *Saccharomyces cerevisiae*

1 Introduction

Taxol is one of the best natural anti-tumor drugs on the market. It is considered to be one of the most effective anti-cancer drugs for human beings in the next 20 years (Yang et al., 2020; Huang et al., 2021). The main source of taxol is derived from taxus plants, which is limited by the taxol abundance and the natural resources of taxus plants. On natural conditions, the growth speed of taxus plants is slow, and regeneration ability is poor, thus the extracted taxol has been unable to meet the demand of clinical from taxus plants (Goldspiel, 1997; Htay and Liu, 2005; Miller et al., 2008). At present, the production of taxol mainly as and semi-synthetic raw material using the natural precursor 10-deacetylbaccatin III. However, the acquisition of 10-deacetylbaccatin III has some limitations, such as low extraction efficiency, high production cost and dependence of raw materials on taxus plants

(Cragg, 1998; Baloglu and Kingston, 1999; Frense, 2007). With the help of microbial metabolic engineering and synthetic biology technology, constructing a suitable heterologous biosynthesis system to produce taxol gradually became a research hotspot (Choi et al., 2019; Zhang et al., 2022). In recent years, the biosynthesis of taxadiene, an important intermediate of taxol, has made some progress (Gallego-Jara et al., 2020; Sabzehzari et al., 2020; Hu et al., 2021).

Taxadiene is a diterpenoid compound derived from cyclized geranyl pyrophosphate (GGPP) substrate of taxadiene synthase (TS) (Hefner et al., 1998; Williams et al., 2000; Walker and Croteau, 2001). GGPP is a common precursor produced from MEP or MVA pathway and generated by geranyl diphosphate synthase (GGPPS) (Vranová et al., 2013; Liao et al., 2016; Chatzivasileiou et al., 2019). To date, taxadiene biosynthesis has been realized and modified in different microbial systems (Moser and Pichler, 2019; Navale et al., 2021). In 2010, Ajikumar *et al.* gained the yield of taxadiene of 300 mg/L in shake flask fermentation through engineering multi-module metabolic pathway and optimizing the metabolic balance of taxadiene synthesis. And then the yield of taxadiene was further increased to 1,020 mg/L by modifying fed-batch fermentation (Ajikumar et al., 2010). To realize the biosynthesis of taxol, taxadiene also needs to undergo a series of complex reactions such as hydroxylation, acetylation and epoxidation (Jennewein and Croteau, 2001; Hu et al., 2021). But the complete intimal system and protein post-translational modification system were lack in the prokaryotic system, which severely limits biosynthesis of taxol using prokaryotic system (Bheri et al., 2020; Ramazi and Zahiri, 2021). Thus, in the past decade, there has been no major breakthrough in the heterosynthesis of taxol in *E. coli* (*E. coli*). And the eukaryotic system was regarded as a potential candidate for heterosynthesis of taxol, since the eukaryotic system was suitable for the expression of multiple heterologous P450 hydroxylases and their reductase, which are necessary for taxol synthesis (Hauser and Matthes, 2017; Jia et al., 2022). Therefore, in recent years, eukaryotic systems have been increasingly favored by researchers in taxadiene biosynthesis. Engels *et al.* screened the sources of GGPPS enzyme, engineered the isoprene pathway and optimized the TS gene, achieving the yield of taxadiene of 8.7 mg/L, which laid a foundation for the production of taxol in *Saccharomyces cerevisiae* (*S. cerevisiae*) (Engels et al., 2008). Behnaz *et al.* enhanced the expression of TS by introducing the solubilizing tags to improve taxadiene titre to 57 mg/L at 30°C. Meanwhile, Behnaz *et al.* also found the temperature-dependant phenomenon of TS, according to that, a maximum taxadiene titre of 129 mg/L was obtained by manipulating the fermentation temperature at 20°C (Nowrouzi et al., 2020). In addition, Li *et al.* introduced taxadiene synthase, and taxadiene-5α-hydroxylase in Nicotiana benthamiana through chloroplastic compartmentalized metabolic engineering to produce taxadiene (56.6 μg/g FW) and taxadiene-5α-ol (1.3 μg/g FW). This also was an alternative eukaryotic platform for taxol production (Li et al., 2019). Previous studies have shown that the accumulation of GGPP is abundant enough to supply downstream metabolic pathways for the production of taxadiene and its derivatives in both prokaryotic and eukaryotic systems (Ober, 2010; Zhou et al., 2012). However, the synthesis efficiency of taxadiene in the eukaryotic system still needs to be further improved compared to that in the prokaryotic system. The eukaryotic cells have abundant subcellular organelle structures, which may lead to different localization of the heterologous proteins for complex natural products (Luo et al., 2015; Du and Li, 2021). Thus, we speculate that key enzymes catalysis compartmentalization may be an important factor to limit the taxadiene biosynthesis in eukaryotic system.

Focusing on the problem of enzymes catalysis compartmentalization, many researchers have done a lot of work and made great achievements. Dawid *et al.* truncated the N-terminal mitochondrial location sequences of valine biosynthesis enzymes Ilv2, Ilv5 and Ilv3 to relocate the truncated enzymes into the cytoplasm, resulting in the increase of isobutanol production (Brat et al., 2012). Jiang *et al.* employed synthetic protein scaffolds to colocalize the sequential enzymes of Idi1, GES-Erg20 and IS to eliminate enzymes catalysis compartmentalization, and the citronellol titer was incerased to 8.30 g/L (Jiang et al., 2021). In addition, protein fusion is also an important means to realize target gene relocation and eliminate enzymes catalysis compartmentalization. Ma *et al.* introduced the peroxisome targeting sequence SKL to fusion proteins CrtW-CrtZ, causing fusion proteins CrtW-CrtZ to relocate to suitable organelle of peroxisome, and the astaxanthin titer was increased to 58.7 mg/L (Ma et al., 2021). In addition, Shi *et al.* fused protopanaxadiol synthase (PPDS) to lipid droplet membrane protein Pln1 for closing the spatial distance to substrate dammarenediol-II (DD) accumulated in lipid droplet, leading to the final product ginsenoside Compound K yield of 5 g/L (Shi et al., 2021).

In our previous study, we successfully constructed a high-yield strain of GGPP by screening the metabolic pathway gene and regulating the promoters of critical genes, which provided sufficient precursors for the biosynthesis of taxadiene (Song et al., 2017). Herein, we verified the catalysis compartmentalization of two key exogenous enzymes GGPPS and TS in taxadiene synthesis pathway (Figure 1). Next, we reconstructed the subcellular localization of TS in *S. cerevisiae* to eliminate enzymes (enzyme) catalysis compartmentalization using the methods of N-terminal signal peptide truncation of TS and enzyme fusion of GGPPS-t60TS, and the taxadiene yield was increased by 21% and 54% respectively. Especially, the GGPPS-TS fusion enzyme is more effective. And then, we improved the expression of GGPPS-TS fusion enzyme *via* the multi-copy plasmid, leading to the higher taxadiene titer of 21.8 mg/L at the shake-flask level. Finally, the taxadiene titer was increased to 184.2 mg/L for 144 h on the optimized fed-batch fermentation conditions in 3 L bioreactor, which is approximately increased to 17.8-fold compared with the initial strain yZCL010. This study is a successful example of improving the biosynthesis of complex natural products by overcoming multistep enzyme catalysis compartmentalization in the eukaryotic microbial cell.

2 Materiails and methods

2.1 Strains and culture media

All the yeast strains used in this study were listed in Table 1. *E. coli* Top10 (TransGen Biotech, Beijing, China) was used for

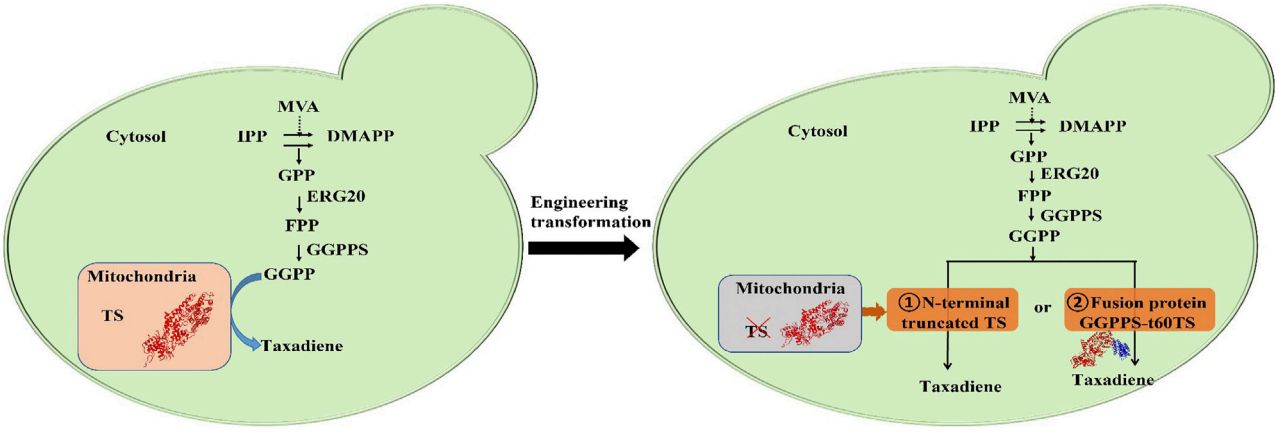


FIGURE 1
The biosynthesis pathway and metabolic engineering transformation of taxadiene. MVA, mevalonate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPPS, geranylgeranyl diphosphate synthase; GGPP, geranylgeranyl diphosphate; TS, taxadiene synthase.

TABLE 1 Yeast strains used in this study.

Strains	Description	Source
BY4742	MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0	Invitrogen
Sc01	Delta: P_{TDH1}-*Sa*GGPPS	This lab
	-P_{GPM1}-BTS1-ERG20-P_{PDC1}-tHMGR; *LEU2*	
Sc02	Delta: P_{TDH1}-*Sa*GGPPS	This lab
	-P_{GPM1}-BTS1-ERG20-P_{PDC1}-tHMGR; *LEU2; Ura3*	
yZCL004	Sc02; pZCL060; *Ura3*	This work
yZCL005	Sc02; pZCL055; *Ura3*	This work
yZCL006	Sc02; pZCL061; *Ura3*	This work
yZCL007	Sc02; pZCL062; *Ura3*	This work
yZCL010	Sc01; pZCL053	This work
yZCL011	Sc01; pZCL063	This work
yZCL012	Sc01; pZCL064	This work
yZCL047	Sc02; pZCL069; *Ura3*	This work
yZCL059	Sc01; pZCL081	This work
yZCL060	Sc01; pZCL082	This work
yZCL061	Sc01; pZCL083	This work

plasmids construction and amplification. BY4742 used for construction of taxadiene producing strains (listed in Table 1), was obtained from EUROSCARF (Frankfurt, Germany) (Entian and Kotter, 2007).

LB medium (0.5% yeast extract, 1% tryptone and 1% NaCl) was used to culture *E. coli* Top 10 for the propagation of recombinant plasmids, supplemented with 50 μg/mL kanamycin or 100 μg/mL ampicillin at 37°C. Yeast cells were routinely cultured in YPD medium (1% yeast extract, 2% peptone, 2% glucose) or synthetic complete (SC) drop-out medium (2% dextrose, 0.67% yeast nitrogen base, 0.2% amino acid mix lacking selected amino acids) at 30°C. The selected plasmids carrying nutritional screening tags were introduced into *S. cerevisiae* cells and cultured in a synthetic complete medium without corresponding amino acids. All the medium formulations for yeast culture was the same as our previous work (Zeng et al., 2020).

2.2 Plasmid construction

All the plasmids are summarized in Supplementary Table S1, and the primers used during plasmid construction are summarized in Supplementary Table S2. The yeast expression plasmid pRS415 and pRS425 were purchased from ADDGENE (American) and the ampicillin resistant gene of the plasmid was substituted by kanamycin resistant gene to construct pRS415K and pRS425K. *Sa*GGPPS gene (ACCESSION P39464) and *Tb*TS gene (ACCESSION Q41594) were codon optimized for expression in *S. cerevisiae* and synthesized by GenScript,Inc. (China) (Supplementary Table S3). The basic gene expression cassette was assembled by overlap extension PCR (OE-PCR) (pZCL026; pZCL027; pZCL041) (Supplementary Figure S2). The OE-PCR products were digested by *BamH* I/*Not* I and inserted into the corresponding sites of expression plasmid pRS415k, obtaining a series of plasmids for the construction of target enzymes localization strains. Then the plasmid with N-terminal truncated *Tb*TS or fusion protein GGPPS-TS (including pZCL063; pZCL064; pZCL071) were constructed based on pZCL041(Supplementary Table S1). The plasmids were transformed of into yeast strains by the LiAc/ssDNA carrier DNA/PEG3350 method. The *S. cerevisiae* strains constructed in this study were listed in Table 1(Gietz and Woods, 2006).

2.3 Fermentation conditions

To prepare seed vials, single isolates of each strain from agar plates were grown for SD Agar plate at 30°C and 200 rpm. And the

preculture was transferred to a 250 mL flask filled with 50 mL fresh culture medium, the initial OD_{600} was 0.2. And then preculture was transferred into 250 mL flasks with 50 mL the same fresh medium with initial OD_{600} of 0.2 and cultivated at 30°C, 250 rpm for 10–12 h until the exponential phase. Finally, the third-grade seeds were transferred to a shake flask and fermented for 120 h. The cultured tertiary seeds were then transferred to 200 mL fresh SC medium for 10 h until OD_{600} reached about 6.0. Then the seed culture was transferred to a bioreactor containing a 2.7 L fermentation medium. The initial fermenter medium was YPD containing 20 g/L glucose. The 10% (v/v) n-dodecane (Sigma-Aldrich) was added to the culture at the beginning of the fermentation to enrich taxadiene, which could minimize the loss of taxadiene the product and protect the cells from phase the toxicity brought by taxadiene (Brennan et al., 2012).

For fed-batch fermentation, the inoculation amount in the fermenter is 10% (v/v), and the seed preparation conditions are the same as those in the shaking flask fermentation. The stirring speed is 300 rpm. The temperature was controlled at 30°C. Air flow is used to supply oxygen to the fermentation tank at 2 vvm, and the pH value in the fermentor was controlled at 6.0 with 3 M NaOH. In addition, the biomass in fermentation and the accumulation of taxadiene in a later stage were increased by nitrogen supplemental. In addition, 5 g/L nitrogen yeast extract was added three times at the initial stage of culture every 7 h during the beginning 48 h of cultivation. After 48 h of fermentation, the temperature dropped to 20°C, and 20% (v/v) dodecane was added for two-phase extractive fermentation, and Galactose inducer was added until the final concentration was 20 g/L. Cell growth and glucose concentration were continuously monitored during fermentation. The cell growth and the glucose concentration were constantly monitored during the fermentation process. 500 g/L glucose was fed periodically into the fermentation to keep the glucose concentration under 1.0 g/L. And the organic layer was harvested for taxadiene analysis by centrifugation of the fermentation broth at 12,000 rpm for 10 min. The GGOH (98% purity) was prepared to construct a standard curve to determine the GGOH yield.

2.4 Analytical methods

GC-TOF/MS analyzed taxadiene and GGOH in fermentation products. The target product extracted from n-dodecane was diluted with n-hexane, then 1 μL sample was injected into Shimazu GC-2030 using a Shimazu GCMS-QP2020 automatic sampler. The sample was detected using a quartz capillary column (30 m × 0.25 mm, 0.25 mm DB-5MS, J&W Scientific, Folsom). Design the relevant parameters of the sample detection method. The injector temperature was set at 260°C. The column effluents were introduced into the ion source (250°C) of TOF/MS. And ions were generated by 40 mA ionization current of a 70 eV electron beam. The mass scan range was 50–800 m/z.

For GC-TOF/MS analysis of taxadiene and GGOH, the column chamber temperature was first kept constant at 70°C for 1 min, then increased to 200°C at a rate of 30°C/min for 1 min. Next it increased to 265°C at a rate of 12°C/min and kept for 3 min. The total run time was 14.75 min. Taxadiene was identified by mass fragments 109 m/z and 122 m/z, and the peak time was 10.32 min. GGOH was identified by mass fragments of 69 m/z, 93 m/z, and 119 m/z, and the peak time was 11.35 min.

2.5 Assay of protein subcellular localization

Fluorescence microscopy was used to observe the distribution of fluorescent signals of target proteins in the subcellular to determine the subcellular localization of target proteins. A single colony of inverters extracted from an SC-U-L agar plate was first cultured into a 20 mL tube containing 5 mL SC-U-L medium and cultured at 30°C and 220 rpm for 20–24 h to achieve the exponential phase. The pre-cultured cells were then transferred to 20 mL test tubes and 5 mL identical fresh medium, initially OD_{600}, and cultured at 30°C and 220 rpm for 48 h. Fluorescence image was observed under a fluorescence microscope and treated with fluorescence microscopy software FCSnap.

3 Results and Discussion

3.1 Subcellular localization of GGPPS and TS in taxadiene biosynthesis pathway

In this study, Geranylgeranyl diphosphate synthase from *Sulfolobus acidocaldarius* (*Sa*GGPPS) and sequence-optimized TS from *Taxus brevifolia* (*Tb* t60TS) was introduced into the high-yield GGPP strain to construct the initial taxadiene production strain yZCL010 (Figure 2A). And the taxadiene titer of 10.2 mg/L and the geranylgeraniol (GGOH) titer of 214.2 mg/L were detected in the production strain yZCL010 (Figure 2B). According to the large accumulation of by-product GGOH in the synthesis of taxadiene, we speculated that the synthesis process of taxadiene with key enzymes of GGPPS and TS is limited, which should be the main reason for the low taxadiene production.

Whereafter, the subcellular localizations of two key enzymes *Tb*t60TS and *Sa*GGPPS for the taxadiene synthesis were detected. We firstly added red fluorescent protein (RFP) tag to the C-terminus of each target enzyme, and co-expressed them with fluorescent protein labeled endogenous subcellular localization protein, such as COXIII-GFP for mitochondrial localization characterization and sole expressed GFP for cytoplasmic localization characterization (Rodrigues et al., 2001; Naithani et al., 2003). Fluorescent microscopic image analysis of RFP and GFP signals showed that *Sa*GGPPS was located in cytoplasm, while t60TS was located in mitochondria (Figure 2C). These results indicate that GGPPS and t60TS localized at different subcellular, which would result in the problem of catalytic compartmentalization between the two sequent enzymes GGPPS and t60TS. Therefore the large amount of GGPP generated in the cytoplasm was difficult to contact with TS to produce taxadiene, due to GGPP hardly is transported through the plastid membrane with significant efficiency (Bick and Lange, 2003; Vranová et al., 2012). Instead, accumulated GGPP was converted into by-product GGOH. In our constructed the initial taxadiene production strain yZCL010, only less than 5% of the precursor GGPP was converted to the target product taxadiene, while almost all the rest was produced into by-product GGOH, which is a great metabolic drain. According to the above key

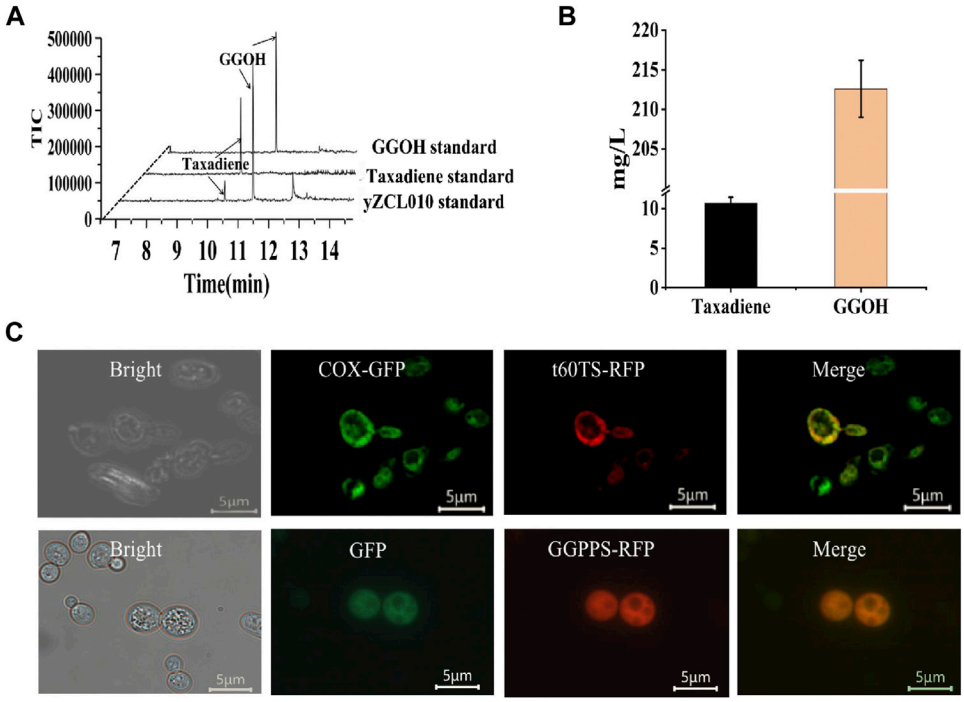


FIGURE 2
Construction of the initial taxadiene production strain and subcellular localization of key exogenous enzymes. **(A)** GC-MS results for the initial strain for taxadiene and GGOH production. The taxadiene and GGOH eluted at about 10.3 min and 11.3 min. **(B)** The titers of taxadiene and GGOH in strain yZCL010. **(C)** Subcellular localization of t60TS and GGPPS in *S. cerevisiae*.

enzymes localization analysis and product detection results, the catalytic compartmentalization of enzymes would reduce the availability of substrates and intermediates, leading to large accumulation of by-product and low yield of the target product.

3.2 Subcellular relocation of TS *via* N-terminal signal peptide truncation

From the above study, *Sa*GGPPS was located in the cytoplasm, which may lead to the accumulation of GGPP in the cytoplasm. Thus, *Tb*t60TS with mitochondria localization will hardly contact with the substrate GGPP to effectively generate taxadiene. Therefore, we attempted to relocate TS from mitochondria into the cytoplasm to improve conversion from GGPP to taxadiene. Here, the proper truncation position for *Tb*TS overexpressed in *S. cerevisiae* was attempted to relocate TS. Based on the predicted results on the secondary structure by https://bioinf.cs.ucl.ac.uk/psipred/ (Supplementary Figure S1A) (McGuffin et al., 2000), N-terminus of *Tb*TS was truncated at other two different positions (R84 and N97) besides M60 (t60TS), according to a certain gradient with slightly change, named t84TS and t97TS respectively (Supplementary Figure S1B).

To determine the subcellular localization of the t84TS and t97TS in *S. cerevisiae*, RFP tag was applied to N-terminal truncated *Tb*TS and cytoplasmic localization protein GFP were also co-expressed (Refer to Section 3.1). Fluorescent microscopic image analysis showed that t84TS was located in the cytoplasm successfully (Figure 3A). In addition, t97TS were also located in the cytoplasm (data not shown).

Meanwhile, N-terminal truncated *Tb*TS (t84TS and t97TS) was introduced into the high-yield GGPP strain to construct the taxadiene production strain of yZCL011, and yZCL012, respectively. As shown in Figure 3B, notably t84TS and t97TS both led to an over 2.5-fold increase in protein soluble expression compared with that of t60TS, in view of the different fluorescence intensity detection. However, the titers of taxadiene only were slightly increased by 20.5% to 12.3 mg/L and by 14.3% to 11.7 mg/L in the production strains of yZCL011 and yZCL012 (Figure 3C), which indicated the taxadiene yield was not correspondingly increased with the enhanced expression of TS. In the case, we further investigated the effect of TS truncation on unit enzyme activity. It is finding that the enzyme activities of t84TS and t97TS with N-terminal deep truncation were either significantly decreased, even less than 50% (Figure 3D). Thus, the deep N-terminal truncation is an undesired strategy for optimizing the function of TS enzyme, even though *Tb*TS was relocate and increased on soluble expression.

3.3 Eliminate enzyme catalysis compartmentalization *via* protein fusion of GGPPS and TS

Since the catalysis compartmentalization between GGPPS and TS is the limiting step for taxadiene synthesis in *S.*

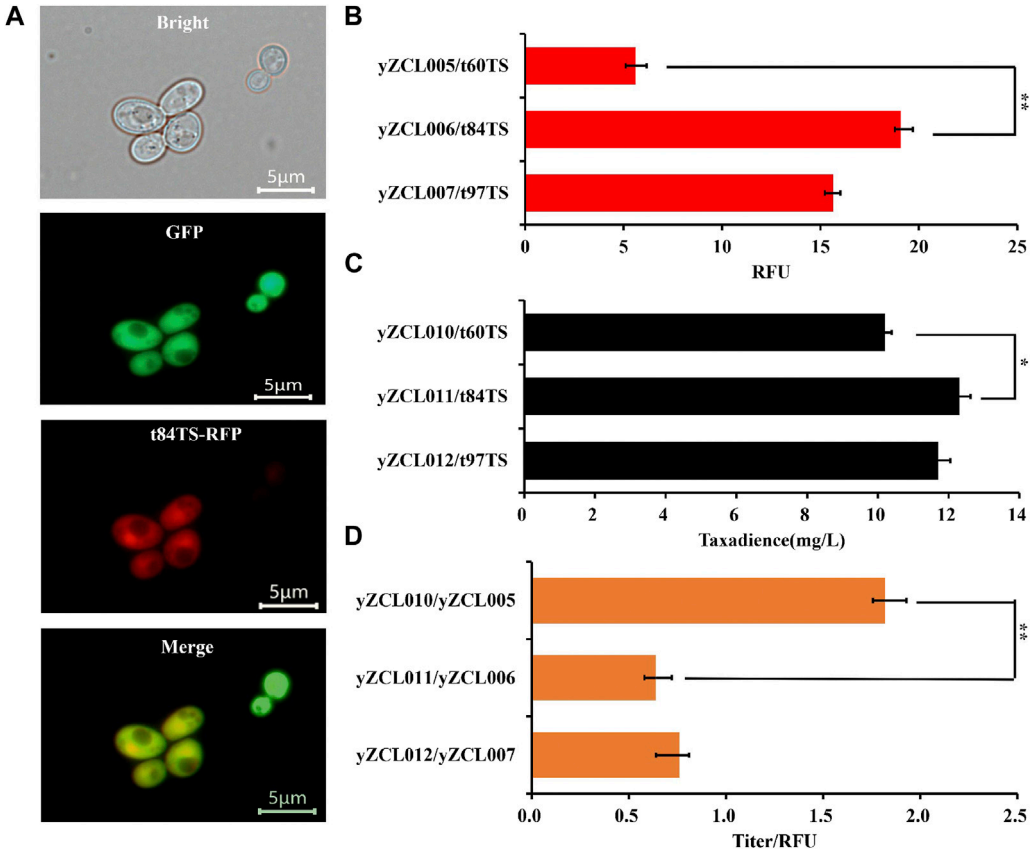


FIGURE 3
Effect of N-terminal truncation on subcellular localization and performance of TS in *S. cerevisiae*. **(A)** Subcellular localization of t84TS in *S. cerevisiae*. **(B)** Effect of N-terminal truncation on the expression of TS. **(C)** Effect of N-terminal truncation of TS on the yield of taxadiene. **(D)** Effect of N-terminal truncation on unit enzyme activity of TS. The error bars represent the means ± SD from three biological replicates. $^{**}p < 0.01$, $^{*}p < 0.05$, student's *t*-test.

cerevisiae, fusion expression of GGPPS and TS would overcome the problem of catalysis compartmentalization to improve sequential catalysis efficiency of GGPPS and TS. Here, the forward and reverse fusion of GGPPS and TS were adopted, and these two proteins were fused with the short flexible linker GSG (Ajikumar et al., 2010). Among them, in the forward fusion, *Sa*GGPPS pulled *Tbt*60TS enzyme to localize in the cytoplasm of *S. cerevisiae* (Figure 4A). Conversely, *Tbt*60TS would traction *Sa*GGPPS to localize in the mitochondria of *S. cerevisiae* in the reverse fusion.

Next, the forward and reverse fusion protein of GGPPS-t60TS and t60TS-GGPPS were introduced into the high-yield GGPP strain to construct the taxadiene production strain yZCL059 and yZCL060, respectively. In the two taxadiene production strains, the yields of taxadiene have been signally improved, reaching to 15.7 mg/L and 13.8 mg/L respectively, relying on the introduction of forward and reverse fusion proteins of GGPPS-t60TS and t60TS-GGPPS (Supplementary Figure S3). Moreover, applying the forward fusion protein of GGPPS-t60TS achieved more significant effect in enhancing taxadiene biosynthesis, and the taxadiene titer of strain yZCL059 was increased by 54%, compared with that of the initial strain yZCL010. To sum up, the forward and reverse fusion protein of GGPPS-t60TS and t60TS-GGPPS were located in cytoplasm and mitochondria of *S. cerevisiae* respectively, and the fusion protein of GGPPS-t60TS in cytoplasm had a greater advantage. It suggested that the cytoplasmic microenvironment seemed be more conducive to the synthesis pathway of taxadiene than mitochondria microenvironment, due to the membrane barrier effect of mitochondria restricting the transfer of intermediates. Moreover, the upstream MVA and FPP synthesis pathways are mainly distributed in the cytoplasm (Vranová E et al., 2013), which will Provide sufficient precursors for the synthesis of GGPP and taxadiene. This may also partly explain the reason of the high production of taxadiene in prokaryotes at present.

In addition, the soluble expression of GGPPS-t60TS was enhanced by about 60% than that of sole t60TS (Figure 4B), which indicates that GGPPS also contributed to the soluble expression of TS, similar to the soluble protein tag effect, in the forward fusion protein. This is a desirable outcome, since the heterologous expression of TS has not been satisfactory in the present study. As well, enzymes fusion in sequential catalysis progress (GGPPS and TS in this study) might minimize the distance between them for higher catalytic activity through enhancing GGPP accessibility for TS (Albertsen et al., 2011). According to the above two favorable effects, the GGPPS-t60TS

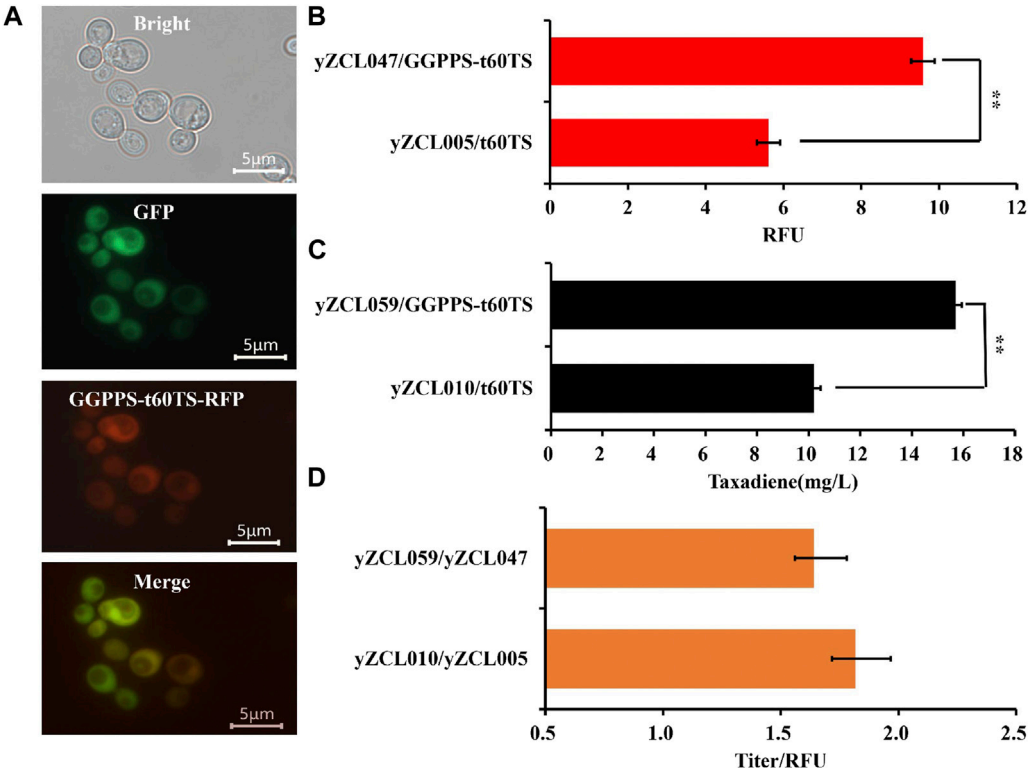


FIGURE 4
Effect of fusion protein GGPPS-t60TS on subcellular localization and performance of TS in *S. cerevisiae*. **(A)** Subcellular localization of fusion protein GGPPS-t60TS in *S. cerevisiae*. **(B)** Effect of fusion protein GGPPS-t60TS on the expression of TS. **(C)** Effect of fusion protein GGPPS-t60TS on the yield of taxadiene. **(D)** Effect of fusion protein GGPPS-t60TS on unit enzyme activity of TS. The error bars represent the means ± SD from three biological replicates. $^{**}p < 0.01$, $^{*}p < 0.05$, student's *t*-test.

raised the maximum yield of taxadiene of 15.7 mg/L in strain yZCL059 (Figure 4C). Meanwhile, there was no significant difference in unit enzyme activity between the fusion protein of GGPPS-t60TS and the initial t60TS (Figure 4D), which takes into account the intrinsic low enzyme activity of TS (Soliman and Tang, 2015). Therefore, the forward fusion protein of GGPPS-t60TS demonstrated the significant advantages in the efficient synthesis of taxadiene in the eukaryotic system.

3.4 Effect of increasing gene copy number on the yield of taxadiene

The forward fusion protein GGPPS-t60TS was proved to be more benefit for taxadiene production. According to the poor soluble expression of TS in yeast on previous studies (Nowrouzi et al., 2020), the expression of the fusion protein GGPPS-t60TS module was enhanced using a multi-copy plasmid to improve the conversion rate of substrate GGPP to taxadiene. We firstly inserted the gene of fusion protein GGPPS-t60TS into the multi-copy plasmid PRS425K at restriction enzymes sites of *BamH* I and *Not* I, and then, the constructed multi-copy plasmid PRS425K -GGPPS-t60TS was introduced into the high-yield GGPP strain to obtained the taxadiene production strain yZCL061. As shown in Figure 5, the taxadiene yield of stain yZCL061was increased by 38.9% than that of stain yZCL059, benefiting from protein expression improvement by multiple-copy plasmid, and taxadiene titer reached to 21.8 mg/L in the shake flask. On the contrary, the by-product GGOH yield was also increased by 4% in the stain yZCL061 compared with that in the stain yZCL059, which did not show statistical significance. Ultimately, the titer ratio of taxadiene to GGOH is enhanced by over 108%, from the titer ratio of 4.7% in initial production strain yZCL010 to 9.8% in stain yZCL061 (Figure 2B, Figure 5). Therefore, the fusion protein of GGPPS-t60TS significantly promoted the conversion efficiency of target product taxadiene.

3.5 Taxadiene overproduction in fed-batch fermentation

To evaluate the performance of the best taxadiene biosynthesis strain (yZCL061), fed-batch fermentations were performed in a 5 L bioreactor using YPD as the batch medium. As shown in Figure 6, a total titer of 184.2 mg/L taxadiene with the maximal biomass at OD_{600} = 110.3 was achieved, which is the highest reported titer in eukaryotic cells. The fermentation process was divided into two stages. The first stage was the cell growth stage, in which the

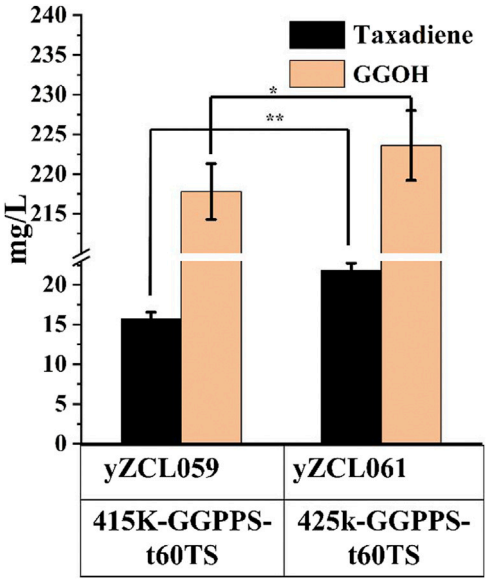


FIGURE 5
The effect of overexpression of fusion protein GGPPS-t60TS on the titers of taxadiene and GGOH. 415K incited single-copy plasmid, 425K incited multi-copy plasmid. The error bars represent the means ± SD from three biological replicates. $^{**}p < 0.01$, $^{*}p < 0.05$, student's *t*-test.

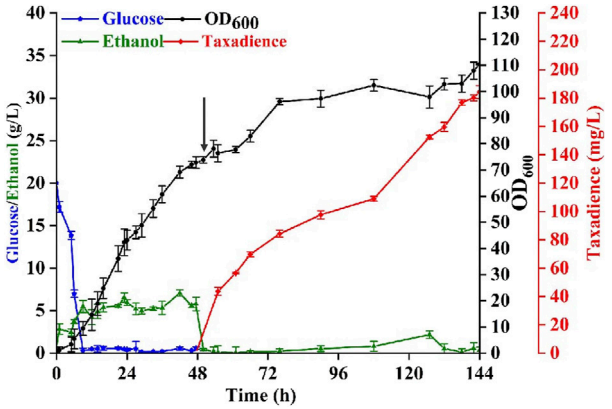


FIGURE 6
Taxadiene production in fed-batch fermentation. Fed-batch fermentations were performed at a 3.0 L scale, using YPD medium and by the engineered *S. cerevisiae* strain yZCL061. Black arrows represented the addition of n-dodecane and galactose into the medium. The error bars represent the means ± SD from three biological replicates.

cultivation temperature was controlled at 30°C for 48 h until the cell density OD_{600} reached 72.8 (Figure 6). The initial glucose concentration was set as 20 g/L. The supplemental carbon source glucose was strictly controlled below 1 g/L due to the carbon source restriction strategy. The second stage taxadiene producing stage was initialized after the temperature was decreased to 20°C and 20 g/L galactose was added at 48 h (Nowrouzi et al., 2020). The supplemental carbon source was changed into ethanol because the existence of glucose would inhibit the expression of the GAL promoters. With fermentation proceeding, the taxadiene production also increased steadily. After 144 h cultivation, the taxadiene titer reached 184.2 ± 0.56 mg/L (Figure 6), which was eight times more than the output at the shake flask level. Even though the titer is still lower than that in *E. coli* (Ajikumar et al., 2010), increasing the yeast cell tolerance to taxadiene would be an efficient solution to minimize the taxadiene production gap between *E. coli* and yeast.

4 Conclusion

In this work, it is found that the catalysis compartmentalization between two key enzymes of GGPPS and TS limited taxadiene production. Applying the intracellular relocation strategies for TS eliminated catalysis compartmentalization *via* N-terminal truncation of TS and enzyme fusion of GGPPS-TS. With the help of forward fusion protein of GGPPS-t60TS, the taxadiene yield was increased by 54% to 15.7 mg/L. Subsequently, the taxadiene titer was further enhanced to 21.8 mg/L at shake-flask level by over-expressing fusion enzyme of GGPPS-t60TS in multi-copy plasmid. Eventually, a highest reported titer of 184.2 mg/L taxadiene in eukaryotic cells was achieved in 3 L fed-batch fermentation. Our study provides an effective strategy to eliminate multistep enzymes catalysis compartmentalization for improving biosynthesis of complex natural products.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

MY and WX conceived the study. CZ and WC participated in strain construction. MY and CZ carried out results data analysis. CZ and TD participated in fed-batch fermentation. CZ and WC carried out the chemical analysis. BL, WX, and YW helped to draft the manuscript. BL and WX participated in design and coordination of the study. MY supervised the whole research and revised the manuscript. All the authors read and approved the final manuscript.

Funding

This work was supported by the Key-Area Research and Development Program of Guangdong Province (2020B0303070002), the National Key Research and Development Program of China (2018YFA0900702) and the National Natural Science Foundation of China (21621004 and 21676192).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1141272/full#supplementary-material

References

Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., et al. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* 330 (6000), 70–74. doi:10.1126/science.1191652

Albertsen, L., Chen, Y., Bach, L. S., Rattleff, S., Maury, J., Brix, S., et al. (2011). Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Appl. Environ. Microbiol.* 77 (3), 1033–1040. doi:10.1128/AEM.01361-10

Baloglu, E., and Kingston, D. G. (1999). A new semisynthesis of paclitaxel from baccatin III. *J. Nat. Prod.* 62 (7), 1068–1071. doi:10.1021/np990040k

Bheri, M., Mahiwal, S., Sanyal, S. K., and Pandey, G. K. (2020). Plant protein phosphatases: What do we know about their mechanism of action? *FEBS J.* 288 (3), 756–785. doi:10.1111/febs.15454

Bick, J. A., and Lange, B. M. (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: Unidirectional transport of intermediates across the chloroplast envelope membrane. *Archives Biochem. biophysics* 415 (2), 146–154. doi:10.1016/s0003-9861(03)00233-9

Brat, D., Weber, C., Lorenzen, W., Bode, H. B., and Boles, E. (2012). Cytosolic re-localization and optimization of valine synthesis and catabolism enables increased isobutanol production with the yeast *Saccharomyces cerevisiae*. *Biotechnol. biofuels* 5 (1), 65–16. doi:10.1186/1754-6834-5-65

Brennan, T. C., Turner, C. D., Krömer, J. O., and Nielsen, L. K. (2012). Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 109 (10), 2513–2522. doi:10.1002/bit.24536

Chatzivasileiou, A. O., Ward, V., Edgar, S. M., and Stephanopoulos, G. (2019). Two-step pathway for isoprenoid synthesis. *Proc. Natl. Acad. Sci.* 116 (2), 506–511. doi:10.1073/pnas.1812935116

Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., and Lee, S. Y. (2019). Systems metabolic engineering strategies: Integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37 (8), 817–837. doi:10.1016/j.tibtech.2019.01.003

Cragg, G. M. (1998). Paclitaxel (Taxol®): A success story with valuable lessons for natural product drug discovery and development. *Med. Res. Rev.* 18 (5), 315–331. doi:10.1002/(SICI)1098-1128(199809)18:5<315::AID-MED3>3.0.CO;2

Du, L., and Li, S. (2021). Compartmentalized biosynthesis of fungal natural products. *Curr. Opin. Biotechnol.* 69, 128–135. doi:10.1016/j.copbio.2020.12.006

Engels, B., Dahm, P., and Jennewein, S. (2008). Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metab. Eng.* 10 (3-4), 201–206. doi:10.1016/j.ymben.2008.03.001

Entian, K. D., and Kötter, P. (2007). 25 yeast genetic strain and plasmid collections. *Methods Microbiol.* 36, 629–666. doi:10.1016/S0580-9517(06)36025-4

Frense, D. (2007). Taxanes: Perspectives for biotechnological production. *Appl. Microbiol. Biotechnol.* 73 (6), 1233–1240. doi:10.1007/s00253-006-0711-0

Gallego-Jara, J., Lozano-Terol, G., Sola-Martínez, R. A., Cánovas-Díaz, M., and de Diego Puente, T. (2020). A compressive review about taxol®: History and future challenges. *Molecules* 25 (24), 5986. doi:10.3390/molecules25245986

Gietz, R. D., and Woods, R. A. (2006). Yeast transformation by the LiAc/SS Carrier DNA/PEG method. Methods Mol. Biol. *Clift. N.J.)* 313, 107–120. doi:10.1385/1-59259-958-3:107

Goldspiel, B. R. (1997). Clinical overview of the taxanes. *Pharmacotherapy* 17 (52), 110s–125s. doi:10.1002/j.1875-9114.1997.tb03813.x

Hauser, K., and Matthes, J. (2017). Medical students' medication communication skills regarding drug prescription a qualitative analysis of simulated physician-patient consultations. *Eur. J. Clin. Pharmacol.* 73 (4), 429–435. doi:10.1007/s00228-016-2192-0

Hefner, J., Ketchum, R. E., and Croteau, R. (1998). Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase fromtaxus canadensisand assessment of the role of this prenyltransferase in cells induced for taxol production. *Archives Biochem. biophysics* 360 (1), 62–74. doi:10.1006/abbi.1998.0926

Htay, T., and Liu, M. W. (2005). Drug-eluting stent: A review and update. *Vasc. Health Risk Manag.* 1 (4), 263–276. doi:10.2147/vhrm.2005.1.4.263

Hu, Z., Liu, X., Tian, M., Ma, Y., Jin, B., Gao, W., et al. (2021). Recent progress and new perspectives for diterpenoid biosynthesis in medicinal plants. *Med. Res. Rev.* 41 (6), 2971–2997. doi:10.1002/med.21816

Huang, M., Lu, J. J., and Ding, J. (2021). Natural products in cancer therapy: Past, present and future. *Nat. Prod. Bioprospecting* 11 (1), 5–13. doi:10.1007/s13659-020-00293-7

Jennewein, S., and Croteau, R. (2001). Taxol: Biosynthesis, molecular genetics, and biotechnological applications. *Appl. Microbiol. Biotechnol.* 57 (1), 13–19. doi:10.1007/s002530100757

Jia, Q., Brown, R., Köllner, T. G., Fu, J., Chen, X., Wong, G. K., et al. (2022). Origin and early evolution of the plant terpene synthase family. *Proc. Natl. Acad. Sci.* 119 (15), e2100361119. doi:10.1073/pnas.2100361119

Jiang, G., Yao, M., Wang, Y., Xiao, W., and Yuan, Y. (2021). A "push-pull-restrain" strategy to improve citronellol production in *Saccharomyces cerevisiae*. *Metab. Eng.* 66, 51–59. doi:10.1016/j.ymben.2021.03.019

Li, J., Mutanda, I., Wang, K., Yang, L., Wang, J., and Wang, Y. (2019). Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in *Nicotiana benthamiana*. *Nat. Commun.* 10 (1), 4850–4912. doi:10.1038/s41467-019-12879-y

Liao, P., Hemmerlin, A., Bach, T. J., and Chye, M. L. (2016). The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnol. Adv.* 34 (5), 697–713. doi:10.1016/j.biotechadv.2016.03.005

Luo, Y., Li, B. Z., Liu, D., Zhang, L., Chen, Y., Jia, B., et al. (2015). Engineered biosynthesis of natural products in heterologous hosts. *Chem. Soc. Rev.* 44 (15), 5265–5290. doi:10.1039/c5cs00025d

Ma, Y., Li, J., Huang, S., and Stephanopoulos, G. (2021). Targeting pathway expression to subcellular organelles improves astaxanthin synthesis in *Yarrowia lipolytica*. *Metab. Eng.* 68, 152–161. doi:10.1016/j.ymben.2021.10.004

McGuffin, L. J., Bryson, K., and Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* 16 (4), 404–405. doi:10.1093/bioinformatics/16.4.404

Miller, K., Neilan, B., and Sze, D. M. (2008). Development of Taxol and other endophyte produced anti-cancer agents. *Recent Pat. Anti-Cancer Drug Discov.* 3 (1), 14–19. doi:10.2174/157489208783478685

Moser, S., and Pichler, H. (2019). Identifying and engineering the ideal microbial terpenoid production host. *Appl. Microbiol. Biotechnol.* 103 (14), 5501–5516. doi:10.1007/s00253-019-09892-y

Naithani, S., Saracco, S. A., Butler, C. A., and Fox, T. D. (2003). Interactions among*COX1*,*COX2*, and*COX3*mRNA-specific translational activator proteins on the inner surface of the mitochondrial inner membrane of*Saccharomyces cerevisiae*. *Saccharomyces cerevisiae*. *Mol. Biol. Cell.* 14 (1), 324–333. doi:10.1091/mbc.e02-08-0490

Navale, G. R., Dharne, M. S., and Shinde, S. S. (2021). Metabolic engineering and synthetic biology for isoprenoid production in *Escherichia coli* and *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 105 (2), 457–475. doi:10.1007/s00253-020-11040-w

Nowrouzi, B., Li, R. A., Walls, L. E., d'Espaux, L., Malcı, K., Liang, L., et al. (2020). Enhanced production of taxadiene in *Saccharomyces cerevisiae*. *Microb. Cell. factories* 19 (1), 200–212. doi:10.1186/s12934-020-01458-2

Ober, D. (2010). Gene duplications and the time thereafter–examples from plant secondary metabolism. *Plant Biol.* 12 (4), 570–577. doi:10.1111/j.1438-8677.2009.00317.x

Ramazi, S., and Zahiri, J. (2021). Post-translational modifications in proteins: Resources, tools and prediction methods. *Database* 2021, baab012. doi:10.1093/database/baab012

Rodrigues, F., van Hemert, M., Steensma, H. Y., Côrte-Real, M., and Leão, C. (2001). Red fluorescent protein (DsRed) as a reporter in *Saccharomyces cerevisiae*. *J. Bacteriol.* 183 (12), 3791–3794. doi:10.1128/jb.183.12.3791-3794.2001

Sabzehzari, M., Zeinali, M., and Naghavi, M. R. (2020). Alternative sources and metabolic engineering of Taxol: Advances and future perspectives. *Biotechnol. Adv.* 43, 107569. doi:10.1016/j.biotechadv.2020.107569

Shi, Y., Wang, D., Li, R., Huang, L., Dai, Z., and Zhang, X. (2021). Engineering yeast subcellular compartments for increased production of the lipophilic natural products ginsenosides. *Metab. Eng.* 67, 104–111. doi:10.1016/j.ymben.2021.06.002

Soliman, S., and Tang, Y. (2015). Natural and engineered production of taxadiene with taxadiene synthase. *Biotechnol. Bioeng.* 112 (2), 229–235. doi:10.1002/bit.25468

Song, T. Q., Ding, M. Z., Zhai, F., Liu, D., Liu, H., Xiao, W. H., et al. (2017). Engineering *Saccharomyces cerevisiae* for geranylgeraniol overproduction by combinatorial design. *Sci. Rep.* 7 (1), 14991–15011. doi:10.1038/s41598-017-15005-4

Vranová, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* 64 (1), 665–700. doi:10.1146/annurev-arplant-050312-120116

Vranová, E., Coman, D., and Gruissem, W. (2012). Structure and dynamics of the isoprenoid pathway network. *Mol. plant* 5 (2), 318–333. doi:10.1093/mp/sss015

Walker, K., and Croteau, R. (2001). Taxol biosynthetic genes. *Phytochemistry* 58 (1), 1–7. doi:10.1016/s0031-9422(01)00160-1

Williams, D. C., Carroll, B. J., Jin, Q., Rithner, C. D., Lenger, S. R., Floss, H. G., et al. (2000). Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of taxol catalyzed by recombinant taxadiene synthase. *Chem. Biol.* 7 (12), 969–977. doi:10.1016/s1074-5521(00)00046-6

Yang, Y. H., Mao, J. W., and Tan, X. L. (2020). Research progress on the source, production, and anti-cancer mechanisms of paclitaxel. *Chin. J. Nat. Med.* 18 (12), 890–897. doi:10.1016/S1875-5364(20)60032-2

Zeng, B. X., Yao, M. D., Wang, Y., Xiao, W. H., and Yuan, Y. J. (2020). Metabolic engineering of *Saccharomyces cerevisiae* for enhanced dihydroartemisinic acid production. *Front. Bioeng. Biotechnol.* 8, 152. doi:10.3389/fbioe.2020.00152

Zhang, G., Wang, H., Zhang, Z., Verstrepen, K. J., Wang, Q., and Dai, Z. (2022). Metabolic engineering of yarrowia lipolytica for terpenoids production: Advances and perspectives. *Crit. Rev. Biotechnol.* 42 (4), 618–633. doi:10.1080/07388551.2021.1947183

Zhou, Y. J., Gao, W., Rong, Q., Jin, G., Chu, H., Liu, W., et al. (2012). Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J. Am. Chem. Soc.* 134 (6), 3234–3241. doi:10.1021/ja2114486

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 23 February 2023
DOI 10.3389/fbioe.2023.1132244

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Jifeng Yuan,
Xiamen University, China
Zhigang Sui,
Dalian Institute of Chemical Physics (CAS), China

*CORRESPONDENCE
Ying Wang,
ying.wang@tju.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 27 December 2022
ACCEPTED 13 February 2023
PUBLISHED 23 February 2023

CITATION
Chen L, Xiao W, Yao M, Wang Y and Yuan Y (2023), Compartmentalization engineering of yeasts to overcome precursor limitations and cytotoxicity in terpenoid production.
Front. Bioeng. Biotechnol. 11:1132244.
doi: 10.3389/fbioe.2023.1132244

Compartmentalization engineering of yeasts to overcome precursor limitations and cytotoxicity in terpenoid production

Lifei Chen[1], Wenhai Xiao[1,2], Mingdong Yao[1], Ying Wang[1]* and Yingjin Yuan[1]

[1]Frontier Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [2]Georgia Tech Shenzhen Institute, Tianjin University, Shenzhen, China

Metabolic engineering strategies for terpenoid production have mainly focused on bottlenecks in the supply of precursor molecules and cytotoxicity to terpenoids. In recent years, the strategies involving compartmentalization in eukaryotic cells has rapidly developed and have provided several advantages in the supply of precursors, cofactors and a suitable physiochemical environment for product storage. In this review, we provide a comprehensive analysis of organelle compartmentalization for terpenoid production, which can guide the rewiring of subcellular metabolism to make full use of precursors, reduce metabolite toxicity, as well as provide suitable storage capacity and environment. Additionally, the strategies that can enhance the efficiency of a relocated pathway by increasing the number and size of organelles, expanding the cell membrane and targeting metabolic pathways in several organelles are also discussed. Finally, the challenges and future perspectives of this approach for the terpenoid biosynthesis are also discussed.

KEYWORDS

compartmentalization, terpenoids, precursor limitation, cell cytotoxicity, metabolic engineering

Introduction

Terpenoids are a large family of more than 8,000 natural compounds with a wide range of applications, such as pharmaceuticals, agrochemicals, food additives and biofuels. All terpenoids are derived from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), *via* the larger prenyl diphosphate compounds farnesyl diphosphate (FPP), geranyl diphosphate (GPP), or geranylgeranyl diphosphate (GGPP), which represent the universal pool of precursors for terpenoid biosynthesis. According to the number of isoprene units (Zhang and Hong, 2020), terpenoids are classified as monoterpenoids (C10), sesquiterpenoids (C15), diterpenoids (C20), triterpenoids (C30), and tetraterpenoids (carotenoids, C40). In recent years, the rapid development of metabolic engineering and synthetic biology has led to the development of alternative approaches meet the increasing demand for terpenoids (Chen et al., 2018). In eukaryotic cells, common strategies include increasing the supply of acetyl-CoA (Chen et al., 2013), upregulation of the mevalonate

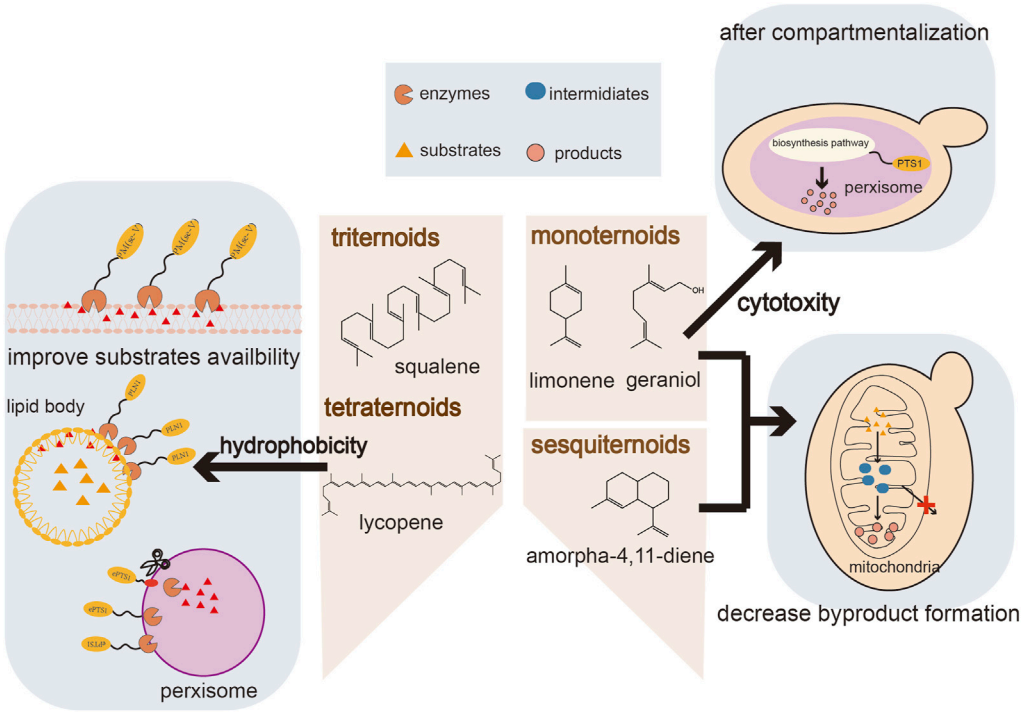


FIGURE 1
An overview of compartmentalization strategies for terpenoid production.

(MVA) pathway (Ro et al., 2006), and engineering that improves the supply of cofactors (Chen R. et al., 2022), some of which are common to all terpenoids (Zhang et al., 2022).

In spite of the common upstream pathway, the enormous structural diversity of terpenoids presents many unique challenges for their biosynthesis, necessitating specific optimization strategies for different types of terpenoids (Figure 1). The insufficient precursor supply is an important factor restricting the production of monoterpenes and sesquiterpenes (Moser and Pichler, 2019). GPP, the precursor of monoterpenes, is also an intermediate of the pathway leading to FPP. Once synthesized, GPP is rapidly converted by FPP synthase (ERG20) which limits the GPP pools and blocks the production of monoterpenoids (Ignea et al., 2014). Likewise, FPP is channeled toward sterol biosynthesis by ERG9, which is essential for cell survival and cannot be completely knocked out. In addition, the cytotoxicity of monoterpenes can have a negative effect on the producing hosts and thereby limits further improvement of monoterpenoid production (Zhu K. et al., 2021). Triterpenes and tetraterpenes are large molecules that accumulate intracellularly in specific subcellular compartments, while some terpenoids are not suitable for intracellular accumulation due to their cytotoxicity. The storge capacity of organelles with hydrophobic environments are essential for the terpenoid production. To overcome these challenges caused by terpenoids, strategies involving the optimization of the expression of key enzymes, tolerance engineering and transporter engineering have been proposed to reduce precursor limitations and cytotoxicity, have resulted in significant improvements in terpenoid production (Peng et al., 2017; Zhang et al., 2017).

Strategies to control the subcellular arrangement of metabolic enzymes are a promising approach for resolving the challenges of terpenoid production mentioned above. Eukaryotic cells contain several specific organelles, including mitochondria, peroxisomes, endoplasmic reticulum (ER), lipid bodies (LDs) and the cell membrane (Dyall et al., 2004). All these organelles possess a complex structure, whereby the specific cofactors, metabolites, and unique physicochemical environments of these organelles offers different conditions for different metabolic pathways (Cao et al., 2020). Some organelles, such as mitochondria, peroxisomes and LDs, have phospholipid membranes, which can increase the local concentration of substrates and enzymes inside the smaller organelle compartments. Compartmentalization also blocks competing reactions and reduces the toxicity of intermediates or substrates by organellar insulation (Gao and Zhou, 2019). The production of terpenoids involves complex metabolic compartmentalization, which partly explains the low yield in terpenoid production achieved to date (Basiony et al., 2022), especially using heterologous yeasts. In heterologous biosynthesis in yeasts, enzymes re-locate to specific organelles and some terpenoids accumulate in membrane-like structures such as LDs. The enzymes and its substates are distributed in different organelles in yeast, which result in the separation of enzyme and substate. (Zhu Z-T. et al., 2021). Targeting the biosynthetic pathway to the same organelle where the desired products are stored enhances the biosynthesis of terpenoids. Subcellular compartmentalization for rewiring of metabolic flux can overcome serveral tackles and enhances the productivity of terpenoids pathways.

Many studies have harnessed subcellular compartments to improve terpenoid biosynthesis in yeasts (Hammer and Avalos,

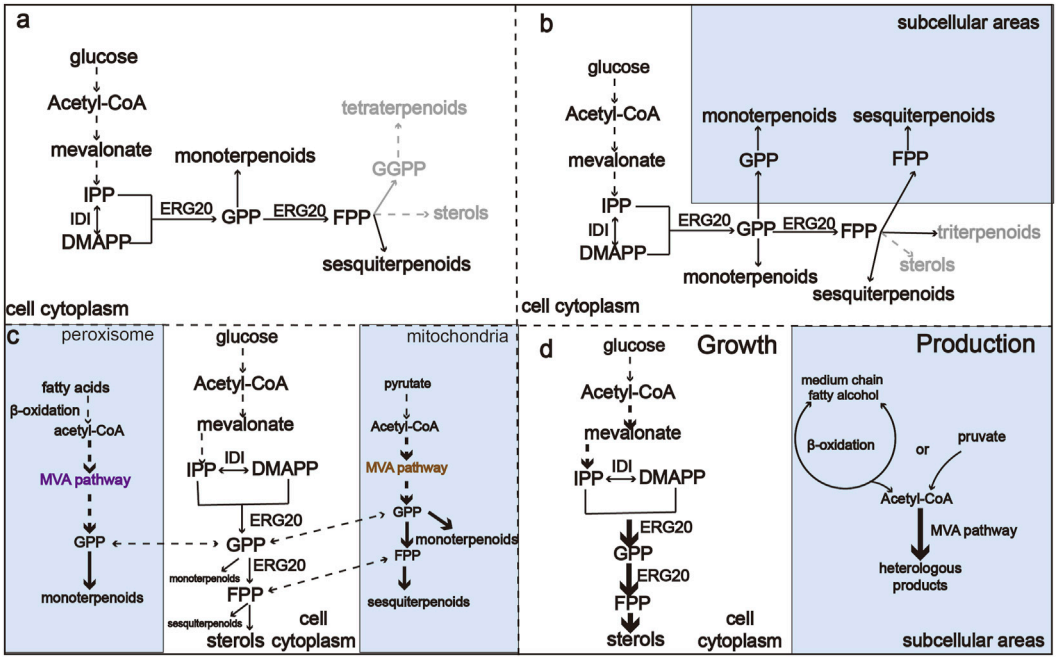


FIGURE 2
An overview of metabolic regulation of monoterpenoid and sesquiterpenoid biosynthesis pathways. **(A)** Monoterpenoid and sesquiterpenoid biosynthesis pathway in the cytoplasm. **(B)** Harnessing GPP and FPP in both cytoplasm and organelles to improve the production of monoterpenoids and sesquiterpenoids. **(C)** Harnessing acetyl-CoA in organelles to improve the production of monoterpenoids and sesquiterpenoids. **(D)** Engineering compartmentalization for orthogonal biosynthesis of monoterpenoids and sesquiterpenoids to reduce competing reactions.

2017; Yocum et al., 2021; Jin et al., 2022), as shown in Table 1. The physiological properties of various organelles as well as their benefits and drawbacks for organelle compartmentalization in terpenoid production (Jin et al., 2022), which will not be discussed below. In this review, we summarize the characteristics of monoterpenoids, sesquiterpenoids and tetraterpenoids and focuse on their unique requirements for improving terpenoid production through organelle compartmentalization. In addition, we provide an overview of recent progress in the modification of organellar morphology for subcellular compartmentalization in terpenoid production.

Monoterpenoids and sesquiterpenoids

Enhancing the supply and utilization of GPP and FPP

The difficulty in creating a sufficient precursor pool of GPP and FPP limits the biosynthesis of all monoterpenoids and sesquiterpenoids (Lei et al., 2021). The traditional monoterpenoid and sesquiterpenoid production processes utilize cytoplasmic GPP and FPP (Figure 2A), which is generally limited. However, additional pools of GPP and FPP are distributed in different organelles. Inspired by the concept that organelles of eukaryotic cells can be harnessed to reconstruct metabolic pathways and thereby increase the availability of precursors (Hammer and Avalos, 2017). Targeting the enzymes or the whole biosynthesis pathway to the organelles can effectively optimize monoterpene and sesquiterpene production (Figure 2B). Optimal utilization of GPP and FPP pools distributed in yeast organelles provides a method for microbial overproduction of monoterpenes and sesquiterpenes (Jia et al., 2020). To harness the pool of FPP in mitochondria, Farhi et al. (2011) targeted FPP synthase and sesquiterpene synthase in mitochondria, achieving an eight- and 20-fold improvements in the production of valencene and amorphadiene, respectively. Upstream of FPP and GPP, acetyl-CoA is the central precursor of the MVA pathway for the biosynthesis of terpenoids, and an insufficient supply of acetyl-CoA limits the metabolic flux toward desired compounds (Krivoruchko et al., 2015). As sites of the β-oxidation of fatty acids, peroxisomes can supply ample acetyl-CoA (van der Klei and Veenhuis, 1997). Similarly, mitochondria have nearly 20–30 times higher acetyl-CoA content than the cytosol (Duran et al., 2020). As shown in Figure 2C, rational organelle compartmentalization can greatly improve the utilization efficiency of the acetyl-CoA pools in different organelles (Son et al., 2022a; Lu et al., 2022). Dong et al. (2021) compartmentalized the whole MVA pathway into mitochondria, leading to a 3.7-fold improvement in the production of α-santalene.

In addition to increasing the supply of precursors to enhance the production of mono- and sesquiterpenes, inhibiting competing pathways is also a major strategy to increase the production of desired compounds (Mai et al., 2021). The FPP-derived squalene and native sterol biosynthesis pathways are the main competing pathways of monoterpenoid and sesquiterpene synthesis. However, these pathways are essential for the fluidity of the yeast cell membrane and host growth (Fischer et al., 2011), which limits heterologous monoterpene and sesquiterpene production (Peng et al., 2017). A commonly used method to reduce the competing

reactions is to construct and screen mutants of ERG20 (Zhang et al., 2020), degrade the rate-limiting enzymes ERG20 and ERG9 (Peng et al., 2018), or establish an orthogonal pathway that uses neryl diphosphate (NPP) as an alternative substrate (Ignea et al., 2019). Due to the necessity of ergosterol for cells, the heterologous pathway affects cellular function due to ubiquitous metabolic interactions. The regulation of metabolic fluxes is crucial to balance growth and production of the desired molecule. However, metabolic rewiring faces many biological challenges that affect the growth and fitness of the hosts (Zhu Z-T. et al., 2021). A method to overcome this obstacle would be to minimize the interactions between heterologous pathway and native metabolism (Pandit et al., 2017; Ignea et al., 2019). This can be achieved by constructing orthogonal pathways through subcellular compartmentalization (Figure 2D). For example, targeting the MVA biosynthesis pathway and the product biosynthesis pathway toorganelles can separate chassis growth from metabolite production. In addition, organelles such as peroxisome and mitochondria can isolate the intermediates from competing reactions, thus increasing the GPP or FPP pool by avoiding consumption by cytoplasmic enzymes (Yuan and Ching, 2016; Yee et al., 2019; Yee et al., 2019; Zhang et al., 2020). By introducing the complete MVA pathway into the peroxisome, Dusséaux et al. (2020) achieve up to a 125-fold increases in the production of geranyl diphosphate-derived compounds compared to the cytosolic pathway. This strategy can be used to produce GPP- or FPP-derived compounds in eukaryotic cells without noticeable effects on strain fitness or viability. Metabolic rewiring has become a promising strategy for enhancing terpenoid production in yeasts.

Reducing the cytotoxicity of monoterpenoids and sesquiterpenoids

The high toxicity of monoterpenoids remains a challenging issue (Erdogan and Ozkan, 2017). Notably, monoterpenoids not only interfere with cell walls and organellar membranes by altering membrane fluidity, structural membrane integrity, and membrane composition, but also induce oxidative stress, and bring more toxic monoterpene hydroperoxides (Brennan et al., 2013; Moon et al., 2020). As shown in Table 2, traditional strategies such as transporter engineering (Hu et al., 2020), *in situ* product extraction (Rolf et al., 2020), and tolerance engineering can improve the tolerance of strains to monoterpenes and reducetoxicity (Brennan et al., 2015; Jezierska and Van Bogaert, 2017; Zhu K. et al., 2021). To address the metabolic toxicity of monoterpenoids, intracellular compartmentalization of their biosynthetic pathways is commonly used to insolate the monoterpenoids into specific organelles (Kulagina et al., 2021). The monoterpene indole alkaloids produced in *Catharanthus roseus* were distributed into several organelles to relieve cell toxicity, which provides new insight into how subcellular compartmentalization can enhance the metabolic flux towards the biosynthesis of target compounds (Courdavault et al., 2014; Guirimand et al., 2020). Peroxisomes are not essential for cell growth (Sibirny, 2016), and are also detoxifying organelles (Breitling et al., 2002). Accordingly, the localization of biosynthetic enzymes or whole biosynthetic pathways to peroxisomes may insulate the toxic components away from the host cell cytosol. The peroxisome can be used as a detoxifying organelle for the synthesis of monoterpenes, including limonene, geraniol, α-pinene, sabinene and camphene (Dusséaux et al., 2020). Dusséaux et al. (2020) constructed peroxisome microfactorties able to accelerate recovery growth by the reducing cytotoxicity of monoterpene products. Gerke et al. (2020) targeted the geraniol biosynthetic enzymes to peroxisomes and improved the geraniol tolerance of the yeast cells to reduce product toxicity, resulting in an 80% increase in the geraniol titer. A combination of traditional metabolic strategies and organelle compartmentalization to alleviate product cytotoxicity may provide new insights.

Triterpenoids and tetraterpenoids

Reducing cytotoxicity of triterpenoids and tetraterpenoids

The long-chain triterpenoids and tetraterpenoids are highly polyunsaturated lipophilic hydrocarbons. Similar to triacylglycerols (TAGs) and sterol esters (SEs), triterpenoids and tetraterpenoids are generally sequestered in specific subcellular areas because of their hydrophobicity (Spanova et al., 2010; Wei et al., 2018). As shown in Figure 3, different organelles such as LDs, peroxisomes, ER and plasma membrane can be used as terpenoid reservoirs due to their similar hydrophobicity (Cunningham and Gantt, 1998; Lee et al., 2009). Cytotoxicity is caused by the excessive accumulation of triterpenoids and tetraterpenoids in the host. The plasma membrane provides storage space for the accumulation of heterologous hydrophobic target compounds (Wu et al., 2017), such as carotenoids (Liu et al., 2016). Targeting the key enzymes to the plasma membrane can effectively alleviate the toxicity of hydrophobic compounds and increase production (Bian et al., 2021; Chen M. et al., 2022). Ye et al. (2018) fused a β-carotene ketolase (CrtW) and β-carotene hydroxylase (CrtZ) and targeted the fusion protein to the cell membrane, which resulted in a 215.4% increase in astaxanthin production. By regulating steady-state fluxes and the availability of intermediate metabolites, compartmentalization engineering in yeasts can significantly shorten the distance between the enzyme and the substrate and increase local concentration of the enzyme and the substrate, thereby improve the product yield (Valm et al., 2017).

Providing a suitable environment for the production of triterpenoids and tetraterpenoids

Organelles such as LDs and peroxisomes can provide a suitable storage environment to relive the cytotoxicity of triterpenoids and tetraterpenoids. LDs and peroxisomes are involved in cellular lipid metabolism, among which peroxisomes play a vital role in α- and β-oxidation of fatty acids. These compartments can provide different intracellular environments for enzyme catalysis, such as adequate substrate supply and cofactor balance (Gao and Zhou, 2019).

TABLE 1 Organelle compartmentalization and optimization of terpenoids production.

	Products	Strategy	Organism	Compartment	Title	Modification of organelle	Reference
monoterpene	Geraniol, R-(+)-limonene, CBGA	• Construct a monoterpenoids producing platform in peroxisome • Insulate the GPP from competing pathway • Used peroxisome as detoxifying organelle	*Saccharomyces cerevisiae*	Peroxisome cytoplasm	5.5 g/L (geraniol) 2.6 g/L	—	Dusséaux et al. (2020)
	Nepetalactol geraniol	• Target geraniol biosynthetic pathway to the mitochondria • Protect the GPP pools from consumption by the cytosolic ergosterol pathway	*Saccharomyces cerevisiae*	mitochondria	227 mg/L[a]	—	Yee et al. (2019)
	Linalool	• Dual mitochondria and cytoplasm engineering in linalool production • Enhance the GPP pools for linalool production	*Saccharomyces cerevisiae*	mitochondria	23.45 mg/L[a]	—	Zhang et al. (2020)
	Isoprene	• Make full use of acetyl-CoA both in cytoplasm and mitochondrial	*Saccharomyces cerevisiae*	Mitochondria cytoplasm	2,527 mg/L	—	Lvet al. (2016)
	Sabinene	• Utilize the mitochondria and cytoplasm GPP pools • Overexpression of mitochondria-related genes to improved sabinene production	*Saccharomyces cerevisiae*	Mitochondria cytoplasm	154.9 mg/L[a]	AIM25, FIS1, LSB3, MBA1	Jia et al. (2020)
	Geraniol	• Target geraniol biosynthetic pathway to peroxisome • Reduce the cell toxicity	*Saccharomyces cerevisiae*	peroxisome	2.75 mg/L[a]	pex30, pex31, atg36	Gerke et al. (2020)
sesquiterpene	Valencene amorphadiene	• Co-locate FDPS and sesquiterpenes synthases to the mitochondria	*Saccharomyces cerevisiae*	mitochondria	1.5 mg/L[a] (Valencene) 20 mg/L[a]	—	Farhi et al. (2011)
	α-santalene	• Reconstruct the whole MVA pathway in mitochondria to harness the precursor pools	*Saccharomyces cerevisiae*	mitochondria	41 mg/L[a]	—	Dong et al. (2021)
	amorpha-4,11-diene	• Harness the mitochondria acetyl-CoA for amorpha-4,11-diene production • reduce loss of FPP to cytosolic competing pathways	*Saccharomyces cerevisiae*	mitochondria	427 mg/L[a]	—	Yuan and Ching (2016)
	α-Humulene	• Utilize the native acetyl-CoA pools in peroxisome	*Yarrowia lipolytica*	peroxisome	3.2 g/L	—	Guo Q et al. (2021)
Triterpene	Squalene	• Dual cytoplasmic-peroxisomal engineering for squalene production	*Saccharomyces cerevisiae*	Peroxisome cytoplasm	11 g/L	—	Liu et al. (2020)
		• Dual MVA pathway in mitochondria and cytoplasm to enhance squalene production	*Saccharomyces cerevisiae*	Mitochondria cytoplasm	21.1 g/L	—	Zhu Z-T et al. (2021)
	ginsenoside	• Target protopanaxadiol synthase (PPDS) to LDs • Increase the volumes of lipid droplets	*Saccharomyces cerevisiae*	lipid droplets	5 g/L	GPD1, PAH1, DGAT1, SEI1	Shi et al. (2021)
	protopanaxadiol	• Construction of the protopanaxadiol pathway in peroxisome • optimization of peroxisome proliferation	*Saccharomyces cerevisiae*	peroxisome	4.1 ± 0.2 mg/L[a]	pex11, pex34, atg36	Choi et al. (2022)
	Squalene protopanaxadiol	• Overexpress ER size regulatory factor to increase the production of squalene and protopanaxadiol	*Saccharomyces cerevisiae*	endoplasmic reticulum	634 ± 11 mg/L[a](squalene)	INO2	Kim et al. (2019)
tetraterpene	Lycopene	• Overexpressed key genes associated with fatty acid synthesis and TAG production and regulate lipid-droplet size to increase lycopene accumulation	*Saccharomyces cerevisiae*	lipid droplets	2.37 g/L	PAH1, DGA1, ACC1, OLE1, FLD1	Ma et al. (2019)

(Continued on following page)

TABLE 1 (*Continued*) Organelle compartmentalization and optimization of terpenoids production.

	Products	Strategy	Organism	Compartment	Title	Modification of organelle	Reference
		• Target lycopene pathway to peroxisome	*Pichia pastoris*	peroxisome	73.9 mg/L[a]	—	Bhataya et al. (2009)
	canthaxanthin	• Introduce the β-carotene ketolase variant OBKTM29 to the plasma membrane	*Saccharomyces cerevisiae*	plasma membrane	1.44 g/L	—	Chen M et al. (2022)
	Lutein	• For enforcing metabolic flux towards α-carotene, re-locatinge-cyclase to the plasma membrane	*Saccharomyces cerevisiae*	plasma membrane	110.4 μg/L	—	Bian et al. (2021)
	astaxanthin	• Target the astaxanthin pathway simultaneously to lipid body, endoplasmic reticulum and peroxisome	*Yarrowia lipolytica*	lipid body, endoplasmic reticulum and peroxisome	858 mg/L	—	Ma et al. (2021)

[a]Fermentation at the shake flask level.

TABLE 2 Comparison of reducing cytotoxicity strategies of monoterpenoids production.

	Strategies	Advantage	Shortage	References
Traditional strategies	Transporter engineering (Isolation of target chemicals from cells through membrane protein)	High transport efficiency; High substrate specificity	Challenges to obtain non-natural transporter by protein engineering	Langevin et al. (2020)
	Extraction of *in situ* products (Transport of inhibitory products into the organic phase)	Simple and effective; Not reliance genetic engineering	The organic phase is not suitable for some water-soluble products	Brennan et al. (2012)
	Tolerance engineering (Modulation of the cellular physiology to tolerate toxic products)	Not rely on genetic engineering; Easy operation	Time consuming and labour-intensive	Li et al. (2021)
Organelle compartmentalization	(Targeting of biosynthesis pathway to organelles)	Increases the concentration of enzymes and substrates; Reduces toxicity effectively and relieves bypass competing pathway monoterpene hydroperoxides	somtimes poor targeting of desired protein into organelle	Grewal et al. (2021)

Peroxisomes are enclosed by a single membrane and are speculated to be an additional cellular storage space for lipophilic products in eukaryotic microbial hosts, acting as an advantageous dynamic depot for storing hydrophobic compounds like squalene (Liu et al., 2020). Liu et al. (2020) compartmentalized the entire squalene synthesis pathway into peroxisomes, followed by dual modulation through cytoplasmic and peroxisomal engineering, which improved the squalene titer to 1.7 g/L. Squalene is an important precursor for triterpenoid biosynthesis, and re-locating triterpene synthases to peroxisomes would be a promising solution for triterpenoid production. Recently, the importance of LDs for the storage of hydrophobic carotenoids was highlighted (Larroude et al., 2018; Bu et al., 2022). LDs serve as specialized platforms for lipid metabolism and storage, with a phospholipid monolayer in which molecules are transported in and out of a LDs (Henne et al., 2020). Ma et al. (2019) observed lycopene accumulated in LDs of *S. cerevisiae*. The hydrophobic compounds are distributed in different parts of LDs due to their different structures (Son et al., 2022b). Adjusting organelle morphology can be a useful strategy for increasing terpenoid production, which will be introduced in detail below. Therefore, the selection of organelles that are naturally lipophilic or hydrophobic may be more conducive to the production of triterpenoids and tetraterpenoids.

Reducing spatial separation between enzyme and product

The triterpenoids and tetraterpenoids always accumulate intracellularly in specific organelles, while enzymes locate in other organelles such as the cytoplasm or ER; thus, the spatial distance between the enzyme and its product may influence the yield of the resulting products. Breaking the barrier between the enzyme and the substrate can reinforce metabolic flux and improve biosynthesis efficiency. The accumulation of sterol metabolites (e.g., zymosterol) in LDs prevents their conversion by enzymes located in ER. Guo et al. reconstituted the post-squalene pathway in LDs and revealed an effective method of pathway compartmentalization (Guo X-J. et al., 2021). The location of enzymes of biosynthesis pathways in membrane-enclosed LDs or peroxisomes can significantly enhance the metabolic flux and promote product titers. Shi et al. (2021) co-localized the pathway enzymes in LDs to eliminate the separation between the substrate dammarenediol-II (DD) and protopanaxadiol synthase (PPDS) of the PPD biosynthesis pathway, which increased the conversion rate of DD to PPD from 17.4% to 86.0%. Compartmentalized reconstitution is an effective way of providing adequate storage and improving the output of triterpenoids and tetraterpenoids.

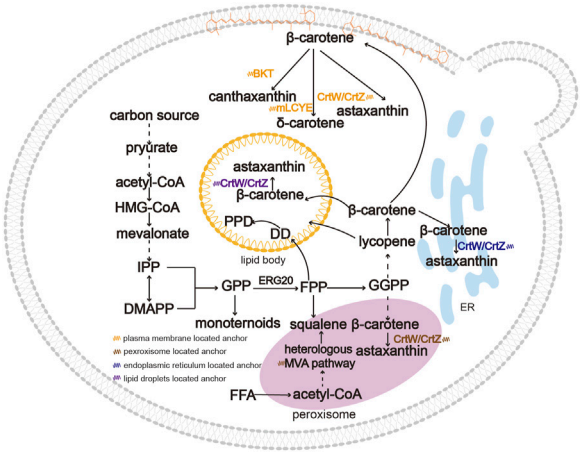


FIGURE 3
An overview of subcellular compartmentalization for the biosynthesis of triterpenoids and tetraterpenoids. Due to their hydrophobicity, triterpenoids and tetraterpenoids are generally sequestered in specific organelles such as LDs, peroxisomes, ER and plasma membrane. Product distribution in strains reduces cytotoxicity and causes spatial separation between substates and products. Moreover, targeting the pathway to more than one compartment can increase the product titer compared to single subcellular compartmentalization. Abbreviations: BKT, β-carotene ketolase; mLCYE, lycopene ε-cyclase.

Modification of organellar morphology

Lipid droplets

Subcellular compartmentalization leads to the co-accumulation and co-localization of enzymes and products in organelles. Son et al. reported that flexible lipids such as squalene with conjugated π bonds are readily dispersed among TAGs, and squalene production can be enhanced by expanding the volume of the LD. Similarly, rigid lipids such as zeaxanthin and β-carotene are retained between TAGs, and the production of zeaxanthin was enhanced by increasing the LD surface area within the cell (Son et al., 2022b). Thus, LDs are suitable organelles for storing triterpenoids and tetraterpenoids, while strategies that control the size and the number of organelles influence their storage capacity (Figure 4). To achieve a high production of lipophilic terpenoids, enhanced the TAG biosynthesis and reduced TAG degradation have been applied to increase LD volume, thereby promoting intracellular terpenoids accumulation (Wei et al., 2018; Ma et al., 2019). Diacylglycerol acyltransferases (DGATs) and phosphatidate phosphatase (PAP) are crucial enzymes that catalyze the ultimate step of TAG synthesis (Adeyo et al., 2011). Overexpression of DGA1 increase lipid content in yeasts, expanding the intracellular storage pool and significantly promoting the accumulation of lycopene (Ma et al., 2019), squalene (Wei et al., 2021) and other terpenoids (Yu et al., 2020). Shi et al. (2021) targeted the limiting enzymes PPDS to LDs and overexpressed DGA1 to increase LD volume, after which the CK titer of the resulting strain reached 5 g/L in 5 L of fed-batch fermentation. Relevant genes that encoding phosphatidic acid phosphatase (PAH1), acetyl-CoA carboxylase (ACC1), and fatty acid desaturase (OLE1) can also regulate TAG biosynthesis (Karanasios et al., 2013; Shi et al., 2014). In addition to LD volume, regulating the number of LDs to increases the monolayer surface is another important strategy affecting terpenoid production (Shi et al., 2021). Overexpressing the *LOA1* gene and deleting *ERD1* can induce the stress response of the ER and stimulate LD formation, which leads to large numbers of LDs (Ayciriex et al., 2012; Teixeira et al., 2018).

Peroxisome

Similar to LDs, alternative methods that manipulate the number and volume of peroxisomes can effectively enhance terpenoid production. By regulating the specific surface area, quantity and size of the organelles, it is possible to enhance the content of its host localization proteins and pathway metabolites, which can further increase the flux of the localized pathway. The morphology of peroxisomes is regulated by peroxins (PEX) and dynamin-related proteins (DRPs), as well as the autophagy (ATG) protein family, which are responsible for peroxisome biogenesis, fission and pexophagy, respectively (Islinger et al., 2018). Most PEX genes are involved in the import of matrix proteins (Heiland and Erdmann, 2005), and the remaining PEX genes are involved in regulation of the abundance and size of peroxisomes (Vizeacoumar et al., 2003; Vizeacoumar et al., 2004; Tower et al., 2011; Yofe et al., 2017). To maximize the storage capacity of the peroxisome membrane of *S. cerevisiae*, Choi et al. increased the expression of *PEX34*, deleted the *PEX11* and *ATG36* gene, after which they introduced a heterologous protopanaxadiol pathway (Choi et al., 2022). Deletion of *PEX11* generated enlarged and clustered peroxisomes, while *PEX34*, overexpression of *PEX34* and deletion of *ATG36* significantly increased the number of peroxisomes. The deletion of *PEX30*, *PEX31*, and *ATG36* in a geraniol producing strains with biosynthetic enzymes targeted to the peroxisome resulted in a larger number of peroxisomes and increased the geraniol titer by 80% (Gerke et al., 2020).

ER and membrane engineering

The expanded membranes of the engineered strains can also be used for additional storage of hydrophobic compounds. Expansion of the ER can increase its capacity to promote the expression of ER-localized proteins and boost terpene accumulation (Kim et al., 2019). Furthermore, suppressing the dephosphorylation and activation of PAH1 can enhance membrane protein productivity and increase the surface area of the ER membrane (Ju et al., 2022). The disruption of PAH1 resulted in a decreased number of LDs, accompanied by the accumulation of neutral lipids in the ER (Adeyo et al., 2011). Similarly to PAH1, overexpression of INO2 in yeast increased its capacity to synthesize endogenous and heterologous ER-associated proteins, leading to a significantly increased squalene titer of 634 mg/L (Kim et al., 2019). Engineering the membrane morphology and improving the membrane synthesis pathway can also increase the stability of membrane-anchored biosynthetic pathways and enhance terpenoid production. A Tsr-augmented recombinant *Escherichia coli* was able to extend the membrane network, enhance the squalene storage and improved the squalene titer to 712 mg/L (Meng et al., 2020). The native *frd* operon and UncF protein also induce membrane invagination and increase the membrane lipid volume (Elmes et al., 1986;

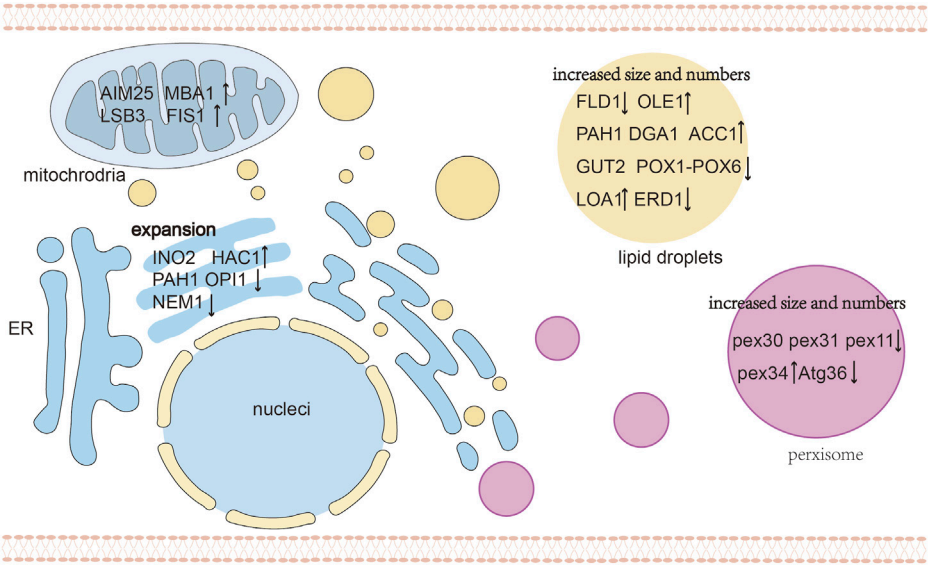


FIGURE 4
An overview of organelle morphology engineering. The target genes of different organelles are shown to regulate organelle morphology, including increased size and number of LDs and peroxisomes, expansion for ER and engineering mitochondrial physiology. Abbreviations: FLD1, seipin; OLE1, fatty acid desaturase; GUT2, G3P dehydrogenase; POX1-6, acyl-CoA oxidases 1 through 6; PEX11, PEX30, PEX32, PEX34, peroxisome-population-regulated proteins; ATG36, autophagy-related protein; OPI1, negative regulator of phospholipid biosynthesis; NEM1, nuclear envelope morphology protein 1; HAC1, Transcriptional activator; ↑ Gene overexpression; ↓ Gene knockdown.

Arechaga et al., 2000). These findings suggest that membrane engineering has achieved good progress in terpenoid production in *E. coli*, and suggest a great potential for the utilization of similar mechanisms to modify the membrane area of yeasts. Membrane engineering is a rapidly developing field with great potential for organelle compartmentalization and metabolic engineering.

Mitochondria

Yeast mitochondria are highly dynamic in size, activity, number, and surface area depending on growth conditions, carbon source and metabolic state (Okamoto and Shaw, 2005). The engineering of mitochondria to increase their activity, morphology, number, and localization can increase the activity of mitochondrial biosynthetic pathways. Jia et al. (2020) overexpressed mitochondria-related proteins including *FIS1* for mitochondrial division, *LSB3* for mitochondrial motility, *MBA1* and *AIM25* encoding other functional proteins to enhance the compartmentalized pathways, and the *AIM25* gene boosted the sabinene titer to 154.9 mg/L. The understanding of mitochondrial physiology and metabolism could be useful to expand the applicability of mitochondrial engineering and realizing its full potential (Ferramosca and Zara, 2021).

Anchoring enzymes simultaneously to several organelles

Organelles such as mitochondria, ER, LDs, and peroxisomes provide different advantages in regulating metabolic fluxes. To further make use of precursors and promote the dynamic balance between different organelles, dual or triple metabolic regulation of the terpenoid synthesis pathway in different subcellular areas could open new possibilities for the production of terpenoids (Lv et al., 2016; Zhu Z-T. et al., 2021; Guo et al., 2022). Comprehensive utilization of multiple organelles can not only enhance the utilization of the limited hydrophobic storage environment, but can also make full use of hydrophobic substrates such as carotene located in different organelles. Ma et al. (2021) expressed the astaxanthin pathway in LDs, endoplasmic reticulum and peroxisomes, and the best strain ultimately produced 858 mg/L of astaxanthin in fed-batch fermentation. Zhu et al. compartmentalized the MVA pathway to mitochondria to improve cell growth and synthesized squalene in the cytoplasm to improve production (Zhu Z-T. et al., 2021). The simultaneous localization of terpenes in different organelles can serve as a promising strategy for optimizing the biosynthetic pathway module and cell growth in parallel.

Conclusion and perspectives

Anchoring the key enzymes or the MVA pathway to specific organelles by localization tags can yield substantial improvements in terpenoid production. In this review, we have systematically analyzed successful approaches using organelle compartmentalization applied to overcome problems associated with the production of different terpenoids, including inadequate supply of GPP/FPP pools and other substrates, cytotoxicity of terpenoids, insufficient storge space, and special challenges of tetraterpenoids. However, more work to improve terpenoid production is still needed in the future. Diterpenoids are one of

the most important families of bioactive compounds, and compartmentalization is still limited in diterpenoid production in yeasts. Li et al. (2019) compartmentalized taxadiene synthase, taxadiene-5ahydroxylase and cytochrome P450 reductase to the chloroplast of *Nicotiana benthamiana* to improve the supply of precursor for taxadiene synthesis. The accumulation of different diterpenoids has temporal and spatial specificity, and compartmentalization of key enzymes and biosynthetic pathway of diterpenoids is a promising alternative to current methods.

As the most extensively studied organelles in cellular physiology, the engineering of mitochondria, peroxisomes, ER, LDs, and plasma membranes has already led to great progress in terpenoid production. However, future studies should target new organelles like the Golgi apparatus and the vacuole whose potential to improve terpenoid production remains unexplored. For secologanin biosynthesis in *C. roseus*, geraniol from the MEP pathway is directed toward the vacuole following a hydroxylation reaction by *G10H* (Verma et al., 2012). Some terpenoids and relevant enzymes are naturally located in the vacuoles (Guirimand et al., 2011), which may provide an alternative option for terpenoid production as a detoxification compartment and proper site for enzyme catalysis. In eukaryotic cells, cofactors are distributed in different subcellular compartments, which necessitates systematically engineering the supply and recycling of cofactors to couple compartmentalized cellular metabolism (Chen R. et al., 2022). Engineering cofactors in organelle compartmentalization benefits the biosynthesis of terpenoids that require cofactors (Cataldo et al., 2020; Zhang et al., 2022). A large number of terpene synthases like P450s and key limiting enzymes such as *HMG1* are NADH-dependent and FADH-dependent enzymes. For instance, CrtI which converts phytoene to lycopene is also a FAD-dependent enzyme (Shen et al., 2016). FAD (H2)-dependent catalytic reactions in the cytosol may be relatively inefficient due to the lower FAD(H) concentration (Pallotta et al., 1998), as the FAD (H2) concentration in mitochondria is 20 times higher than in the cytosol of *S. cerevisiae*. Different cofactors have their unique organelle localization and construction of a cofactor biosynthesis pathway in the cytosol or targeted organelles is an alternative method for increasing terpenoid production.

Author contributions

LC wrote the manuscript. MY, WX, and YY helped to draft the manuscript and revised the manuscript. YW revised and supervised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (32071415).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adeyo, O., Horn, P. J., Lee, S., Binns, D. D., Chandrahas, A., Chapman, K. D., et al. (2011). The yeast lipin orthologue Pah1p is important for biogenesis of lipid droplets. *J. Cell Biol.* 192 (6), 1043–1055. doi:10.1083/jcb.201010111

Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., de Kruijff, B., Runswick, M. J., et al. (2000). Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. *FEBS Lett.* 482 (3), 215–219. doi:10.1016/s0014-5793(00)02054-8

Ayciriex, S., Le Guédard, M., Camougrand, N., Velours, G., Schoene, M., Leone, S., et al. (2012). YPR139c/LOA1 encodes a novel lysophosphatidic acid acyltransferase associated with lipid droplets and involved in TAG homeostasis. *Mol. Biol. Cell* 23 (2), 233–246. doi:10.1091/mbc.E11-07-0650

Basiony, M., Ouyang, L., Wang, D., Yu, J., Zhou, L., Zhu, M., et al. (2022). Optimization of microbial cell factories for astaxanthin production: Biosynthesis and regulations, engineering strategies and fermentation optimization strategies. *Synthetic Syst. Biotechnol.* 7 (2), 689–704. doi:10.1016/j.synbio.2022.01.002

Bhataya, A., Schmidt-Dannert, C., and Lee, P. C. (2009). Metabolic engineering of Pichia pastoris X-33 for lycopene production. *Process Biochem.* 44 (10), 1095–1102. doi:10.1016/j.procbio.2009.05.012

Bian, Q., Zhou, P., Yao, Z., Li, M., Yu, H., and Ye, L. (2021). Heterologous biosynthesis of lutein in *S. cerevisiae* enabled by temporospatial pathway control. *Metab. Eng.* 67, 19–28. doi:10.1016/j.ymben.2021.05.008

Breitling, R., Sharif, O., Hartman, M. L., and Krisans, S. K. (2002). Loss of compartmentalization causes misregulation of lysine biosynthesis in peroxisome-deficient yeast cells. *Eukaryot. Cell* 1 (6), 978–986. doi:10.1128/ec.1.6.978-986.2002

Brennan, T. C., Krömer, J. O., and Nielsen, L. K. (2013). Physiological and transcriptional responses of *Saccharomyces cerevisiae* to d-limonene show changes to the cell wall but not to the plasma membrane. *Appl. Environ. Microbiol.* 79 (12), 3590–3600. doi:10.1128/aem.00463-13

Brennan, T. C., Turner, C. D., Krömer, J. O., and Nielsen, L. K. (2012). Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 109 (10), 2513–2522. doi:10.1002/bit.24536

Brennan, T. C., Williams, T. C., Schulz, B. L., Palfreyman, R. W., Krömer, J. O., and Nielsen, L. K. (2015). Evolutionary engineering improves tolerance for replacement jet fuels in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 81 (10), 3316–3325. doi:10.1128/aem.04144-14

Bu, X., Lin, J. Y., Duan, C. Q., Koffas, M. A. G., and Yan, G. L. (2022). Dual regulation of lipid droplet-triacylglycerol metabolism and ERG9 expression for improved β-carotene production in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 21 (1), 3. doi:10.1186/s12934-021-01723-y

Cao, X., Yang, S., Cao, C., and Zhou, Y. J. (2020). Harnessing sub-organelle metabolism for biosynthesis of isoprenoids in yeast. *Synth. Syst. Biotechnol.* 5 (3), 179–186. doi:10.1016/j.synbio.2020.06.005

Cataldo, V. F., Arenas, N., Salgado, V., Camilo, C., Ibáñez, F., and Agosin, E. (2020). Heterologous production of the epoxycarotenoid violaxanthin in *Saccharomyces cerevisiae*. *Metab. Eng.* 59, 53–63. doi:10.1016/j.ymben.2020.01.006

Chen, M., Li, M., Ye, L., and Yu, H. (2022). Construction of canthaxanthin-producing yeast by combining spatiotemporal regulation and pleiotropic drug resistance engineering. *ACS Synth. Biol.* 11 (1), 325–333. doi:10.1021/acssynbio.1c00437

Chen, R., Gao, J., Yu, W., Chen, X., Zhai, X., Chen, Y., et al. (2022). Engineering cofactor supply and recycling to drive phenolic acid biosynthesis in yeast. *Nat. Chem. Biol.* 18 (5), 520–529. doi:10.1038/s41589-022-01014-6

Chen, X., Gao, C., Guo, L., Hu, G., Luo, Q., Liu, J., et al. (2018). DCEO biotechnology: Tools to design, construct, evaluate, and optimize the metabolic pathway for biosynthesis of chemicals. *Chem. Rev.* 118 (1), 4–72. doi:10.1021/acs.chemrev.6b00804

Chen, Y., Daviet, L., Schalk, M., Siewers, V., and Nielsen, J. (2013). Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metab. Eng.* 15, 48–54. doi:10.1016/j.ymben.2012.11.002

Choi, B. H., Kang, H. J., Kim, S. C., and Lee, P. C. (2022). Organelle engineering in yeast: Enhanced production of protopanaxadiol through manipulation of peroxisome proliferation in *Saccharomyces cerevisiae*. Microorg. *[Online]* 10 (3), 650. doi:10.3390/microorganisms10030650

Courdavault, V., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., and Burlat, V. (2014). A look inside an alkaloid multisite plant: The Catharanthus logistics. *Curr. Opin. Plant Biol.* 19, 43–50. doi:10.1016/j.pbi.2014.03.010

Cunningham, F. X., and Gantt, E. (1998). Genes and enzymes of carotenoid biosynthesis in plants. *Annu. Rev. PLANT PHYSIOLOGY PLANT Mol. Biol.* 49, 557–583. doi:10.1146/annurev.arplant.49.1.557

Dong, C., Shi, Z., Huang, L., Zhao, H., Xu, Z., and Lian, J. (2021). Cloning and characterization of a panel of mitochondrial targeting sequences for compartmentalization engineering in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 118 (11), 4269–4277. doi:10.1002/bit.27896

Duran, L., López, J. M., and Avalos, J. L. (2020). ¡Viva la mitochondria!: Harnessing yeast mitochondria for chemical production. *FEMS Yeast Res.* 20 (6), foaa037. doi:10.1093/femsyr/foaa037

Dusséaux, S., Wajn, W. T., Liu, Y., Ignea, C., and Kampranis, S. C. (2020). Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. *Proc. Natl. Acad. Sci. U. S. A.* 117 (50), 31789–31799. doi:10.1073/pnas.2013968117

Dyall, S. D., Brown, M. T., and Johnson, P. J. (2004). Ancient invasions: From endosymbionts to organelles. *Science* 304 (5668), 253–257. doi:10.1126/science.1094884

Elmes, M. L., Scraba, D. G., and Weiner, J. H. (1986). Isolation and characterization of the tubular organelles induced by fumarate reductase overproduction in *Escherichia coli*. *J. Gen. Microbiol.* 132 (6), 1429–1439. doi:10.1099/00221287-132-6-1429

Erdogan, A., and Ozkan, A. (2017). Investigatıon of antioxıdative, cytotoxic, membrane-damaging and membrane-protective effects of the essential oil of *Origanum majorana* and its oxygenated monoterpene component linalool in human-derived hep G2 cell line. *Iran. J. Pharm. Res.* 16 (1), 24–34.

Farhi, M., Marhevka, E., Masci, T., Marcos, E., Eyal, Y., Ovadis, M., et al. (2011). Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab. Eng.* 13 (5), 474–481. doi:10.1016/j.ymben.2011.05.001

Ferramosca, A., and Zara, V. (2021). Mitochondrial carriers and substrates transport network: A lesson from *Saccharomyces cerevisiae*. *Int. J. Mol. Sci.* 22 (16), 8496. doi:10.3390/ijms22168496

Fischer, M. J., Meyer, S., Claudel, P., Bergdoll, M., and Karst, F. (2011). Metabolic engineering of monoterpene synthesis in yeast. *Biotechnol. Bioeng.* 108 (8), 1883–1892. doi:10.1002/bit.23129

Gao, J., and Zhou, Y. J. (2019). Repurposing peroxisomes for microbial synthesis for biomolecules. *Methods Enzymol.* 617, 83–111. doi:10.1016/bs.mie.2018.12.004

Gerke, J., Frauendorf, H., Schneider, D., Wintergoller, M., Hofmeister, T., Poehlein, A., et al. (2020). Production of the fragrance geraniol in peroxisomes of a product-tolerant baker's yeast. *Front. Bioeng. Biotechnol.* 8, 582052. doi:10.3389/fbioe.2020.582052

Grewal, P. S., Samson, J. A., Baker, J. J., Choi, B., and Dueber, J. E. (2021). Peroxisome compartmentalization of a toxic enzyme improves alkaloid production. *Nat. Chem. Biol.* 17 (1), 96–103. doi:10.1038/s41589-020-00668-4

Guirimand, G., Guihur, A., Ginis, O., Poutrain, P., Héricourt, F., Oudin, A., et al. (2011). The subcellular organization of strictosidine biosynthesis in Catharanthus roseus epidermis highlights several trans-tonoplast translocations of intermediate metabolites. *Febs J.* 278 (5), 749–763. doi:10.1111/j.1742-4658.2010.07994.x

Guirimand, G., Guihur, A., Perello, C., Phillips, M., Mahroug, S., Oudin, A., et al. (2020). Cellular and subcellular compartmentation of the 2C-Methyl-D-Erythritol 4-phosphate pathway in the Madagascar periwinkle. Plants *[Online]* 9 (4), 462. doi:10.3390/plants9040462

Guo, Q., Li, Y. W., Yan, F., Li, K., Wang, Y. T., Ye, C., et al. (2022). Dual cytoplasmic-peroxisomal engineering for high-yield production of sesquiterpene α-humulene in Yarrowia lipolytica. *Biotechnol. Bioeng.* 119 (10), 2819–2830. doi:10.1002/bit.28176

Guo, Q., Shi, T. Q., Peng, Q. Q., Sun, X. M., Ji, X. J., and Huang, H. (2021). Harnessing yarrowia lipolytica peroxisomes as a subcellular factory for α-humulene overproduction. *J. Agric. Food Chem.* 69 (46), 13831–13837. doi:10.1021/acs.jafc.1c05897

Guo, X.-J., Yao, M.-D., Xiao, W.-H., Wang, Y., Zhao, G.-R., and Yuan, Y.-J. (2021). Compartmentalized reconstitution of post-squalene pathway for 7-dehydrocholesterol overproduction in *Saccharomyces cerevisiae*. *Front. Microbiol.* 12, 663973. doi:10.3389/fmicb.2021.663973

Hammer, S. K., and Avalos, J. L. (2017). Harnessing yeast organelles for metabolic engineering. *Nat. Chem. Biol.* 13 (8), 823–832. doi:10.1038/nchembio.2429

Heiland, I., and Erdmann, R. (2005). Biogenesis of peroxisomes. Topogenesis of the peroxisomal membrane and matrix proteins. *Febs J.* 272 (10), 2362–2372. doi:10.1111/j.1742-4658.2005.04690.x

Henne, M., Goodman, J. M., and Hariri, H. (2020). Spatial compartmentalization of lipid droplet biogenesis. *Biochimica Biophysica Acta (BBA) - Mol. Cell Biol. Lipids* 1865 (1), 158499. doi:10.1016/j.bbalip.2019.07.008

Hu, Z., Li, H., Weng, Y., Li, P., Zhang, C., and Xiao, D. (2020). Improve the production of D-limonene by regulating the mevalonate pathway of *Saccharomyces cerevisiae* during alcoholic beverage fermentation. *J. Ind. Microbiol. Biotechnol.* 47 (12), 1083–1097. doi:10.1007/s10295-020-02329-w

Ignea, C., Pontini, M., Maffei, M. E., Makris, A. M., and Kampranis, S. C. (2014). Engineering monoterpene production in yeast using a synthetic dominant negative geranyl diphosphate synthase. *ACS Synth. Biol.* 3 (5), 298–306. doi:10.1021/sb400115e

Ignea, C., Raadam, M. H., Motawia, M. S., Makris, A. M., Vickers, C. E., and Kampranis, S. C. (2019). Orthogonal monoterpenoid biosynthesis in yeast constructed on an isomeric substrate. *Nat. Commun.* 10 (1), 3799. doi:10.1038/s41467-019-11290-x

Islinger, M., Voelkl, A., Fahimi, H. D., and Schrader, M. (2018). The peroxisome: An update on mysteries 2.0. *Histochem Cell Biol.* 150 (5), 443–471. doi:10.1007/s00418-018-1722-5

Jezierska, S., and Van Bogaert, I. N. A. (2017). Crossing boundaries: The importance of cellular membranes in industrial biotechnology. *J. Ind. Microbiol. Biotechnol.* 44 (4-5), 721–733. doi:10.1007/s10295-016-1858-z

Jia, H., Chen, T., Qu, J., Yao, M., Xiao, W., Wang, Y., et al. (2020). Collaborative subcellular compartmentalization to improve GPP utilization and boost sabinene accumulation in *Saccharomyces cerevisiae*. *Biochem. Eng. J.* 164, 107768. doi:10.1016/j.bej.2020.107768

Jin, K., Xia, H., Liu, Y., Li, J., Du, G., Lv, X., et al. (2022). Compartmentalization and transporter engineering strategies for terpenoid synthesis. *Microb. Cell Fact.* 21 (1), 92. doi:10.1186/s12934-022-01819-z

Ju, H., Zhang, C., He, S., Nan, W., and Lu, W. (2022). Construction and optimization of *Saccharomyces cerevisiae* for synthesizing forskolin. *Appl. Microbiol. Biotechnol.* 106 (5), 1933–1944. doi:10.1007/s00253-022-11819-z

Karanasios, E., Barbosa, A. D., Sembongi, H., Mari, M., Han, G. S., Reggiori, F., et al. (2013). Regulation of lipid droplet and membrane biogenesis by the acidic tail of the phosphatidate phosphatase Pah1p. *Mol. Biol. Cell* 24 (13), 2124–2133. doi:10.1091/mbc.E13-01-0021

Kim, J.-E., Jang, I.-S., Son, S.-H., Ko, Y.-J., Cho, B.-K., Kim, S. C., et al. (2019). Tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum for functional assembly of terpene synthesis pathway. *Metab. Eng.* 56, 50–59. doi:10.1016/j.ymben.2019.08.013

Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y., and Nielsen, J. (2015). Microbial acetyl-CoA metabolism and metabolic engineering. *Metab. Eng.* 28, 28–42. doi:10.1016/j.ymben.2014.11.009

Kulagina, N., Besseau, S., Papon, N., and Courdavault, V. (2021). Peroxisomes: A new hub for metabolic engineering in yeast. *Front. Bioeng. Biotechnol.* 9, 659431. doi:10.3389/fbioe.2021.659431

Langevin, A. M., El Meouche, I., and Dunlop, M. J. (2020). Mapping the role of AcrAB-TolC efflux pumps in the evolution of antibiotic resistance reveals near-MIC treatments facilitate resistance acquisition. *mSphere* 5 (6), 010566–e1120. doi:10.1128/mSphere.01056-20

Larroude, M., Celinska, E., Back, A., Thomas, S., Nicaud, J. M., and Ledesma-Amaro, R. (2018). A synthetic biology approach to transform Yarrowia lipolytica into a competitive biotechnological producer of β-carotene. *Biotechnol. Bioeng.* 115 (2), 464–472. doi:10.1002/bit.26473

Lee, P. C., Yoon, Y., and Schmidt-Dannert, C. (2009). Investigation of cellular targeting of carotenoid pathway enzymes in Pichia pastoris. *J. Biotechnol.* 140 (3-4), 227–233. doi:10.1016/j.jbiotec.2009.01.019

Lei, D., Qiu, Z., Qiao, J., and Zhao, G.-R. (2021). Plasticity engineering of plant monoterpene synthases and application for microbial production of monoterpenoids. *Biotechnol. Biofuels* 14 (1), 147. doi:10.1186/s13068-021-01998-8

Li, J., Mutanda, I., Wang, K., Yang, L., Wang, J., and Wang, Y. (2019). Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in Nicotiana benthamiana. *Nat. Commun.* 10 (1), 4850. doi:10.1038/s41467-019-12879-y

Li, J., Zhu, K., Miao, L., Rong, L., Zhao, Y., Li, S., et al. (2021). Simultaneous improvement of limonene production and tolerance in yarrowia lipolytica through tolerance engineering and evolutionary engineering. *ACS Synth. Biol.* 10 (4), 884–896. doi:10.1021/acssynbio.1c00052

Liu, G. S., Li, T., Zhou, W., Jiang, M., Tao, X. Y., Liu, M., et al. (2020). The yeast peroxisome: A dynamic storage depot and subcellular factory for squalene overproduction. *Metab. Eng.* 57, 151–161. doi:10.1016/j.ymben.2019.11.001

Liu, P., Sun, L., Sun, Y., Shang, F., and Yan, G. (2016). Decreased fluidity of cell membranes causes a metal ion deficiency in recombinant *Saccharomyces cerevisiae* producing carotenoids. *J. Ind. Microbiol. Biotechnol.* 43 (4), 525–535. doi:10.1007/s10295-015-1728-0

Lu, S., Zhou, C., Guo, X., Du, Z., Cheng, Y., Wang, Z., et al. (2022). Enhancing fluxes through the mevalonate pathway in *Saccharomyces cerevisiae* by engineering the HMGR and β-alanine metabolism. *Microb. Biotechnol.* 15 (8), 2292–2306. doi:10.1111/1751-7915.14072

Lv, X., Wang, F., Zhou, P., Ye, L., Xie, W., Xu, H., et al. (2016). Dual regulation of cytoplasmic and mitochondrial acetyl-CoA utilization for improved isoprene production in *Saccharomyces cerevisiae*. *Nat. Commun.* 7, 12851. doi:10.1038/ncomms12851

Ma, T., Shi, B., Ye, Z., Li, X., Liu, M., Chen, Y., et al. (2019). Lipid engineering combined with systematic metabolic engineering of *Saccharomyces cerevisiae* for high-yield production of lycopene. *Metab. Eng.* 52, 134–142. doi:10.1016/j.ymben.2018.11.009

Ma, Y., Li, J., Huang, S., and Stephanopoulos, G. (2021). Targeting pathway expression to subcellular organelles improves astaxanthin synthesis in Yarrowia lipolytica. *Metab. Eng.* 68, 152–161. doi:10.1016/j.ymben.2021.10.004

Mai, J., Li, W., Ledesma-Amaro, R., and Ji, X. J. (2021). Engineering plant sesquiterpene synthesis into yeasts: A review. *J. Agric. Food Chem.* 69 (33), 9498–9510. doi:10.1021/acs.jafc.1c03864

Meng, Y., Shao, X., Wang, Y., Li, Y., Zheng, X., Wei, G., et al. (2020). Extension of cell membrane boosting squalene production in the engineered *Escherichia coli*. *Biotechnol. Bioeng.* 117 (11), 3499–3507. doi:10.1002/bit.27511

Moon, J. E., Heo, W., Lee, S. H., Lee, S. H., Lee, H. G., Lee, J. H., et al. (2020). Trehalose protects the probiotic yeast Saccharomyces boulardii against oxidative stress-induced cell death. *J. Microbiol. Biotechnol.* 30 (1), 54–61. doi:10.4014/jmb.1906.06041

Moser, S., and Pichler, H. (2019). Identifying and engineering the ideal microbial terpenoid production host. *Appl. Microbiol. Biotechnol.* 103 (14), 5501–5516. doi:10.1007/s00253-019-09892-y

Okamoto, K., and Shaw, J. M. (2005). Mitochondrial morphology and dynamics in yeast and multicellular eukaryotes. *Annu. Rev. Genet.* 39, 503–536. doi:10.1146/annurev.genet.38.072902.093019

Pallotta, M. L., Brizio, C., Fratianni, A., De Virgilio, C., Barile, M., and Passarella, S. (1998). *Saccharomyces cerevisiae* mitochondria can synthesise FMN and FAD from externally added riboflavin and export them to the extramitochondrial phase. *FEBS Lett.* 428 (3), 245–249. doi:10.1016/s0014-5793(98)00544-4

Pandit, A. V., Srinivasan, S., and Mahadevan, R. (2017). Redesigning metabolism based on orthogonality principles. *Nat. Commun.* 8 (1), 15188. doi:10.1038/ncomms15188

Peng, B., Nielsen, L. K., Kampranis, S. C., and Vickers, C. E. (2018). Engineered protein degradation of farnesyl pyrophosphate synthase is an effective regulatory mechanism to increase monoterpene production in *Saccharomyces cerevisiae*. *Metab. Eng.* 47, 83–93. doi:10.1016/j.ymben.2018.02.005

Peng, B., Plan, M. R., Chrysanthopoulos, P., Hodson, M. P., Nielsen, L. K., and Vickers, C. E. (2017). A squalene synthase protein degradation method for improved sesquiterpene production in *Saccharomyces cerevisiae*. *Metab. Eng.* 39, 209–219. doi:10.1016/j.ymben.2016.12.003

Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440 (7086), 940–943. doi:10.1038/nature04640

Rolf, J., Julsing, M. K., Rosenthal, K., and Lütz, S. (2020). A gram-scale limonene production process with engineered *Escherichia coli*. *Molecules* 25 (8), 1881. doi:10.3390/molecules25081881

Shen, H. J., Cheng, B. Y., Zhang, Y. M., Tang, L., Li, Z., Bu, Y. F., et al. (2016). Dynamic control of the mevalonate pathway expression for improved zeaxanthin production in *Escherichia coli* and comparative proteome analysis. *Metab. Eng.* 38, 180–190. doi:10.1016/j.ymben.2016.07.012

Shi, S., Chen, Y., Siewers, V., and Nielsen, J. (2014). Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1. *mBio* 5 (3), e01130–e01114. doi:10.1128/mBio.01130-14

Shi, Y., Wang, D., Li, R., Huang, L., Dai, Z., and Zhang, X. (2021). Engineering yeast subcellular compartments for increased production of the lipophilic natural products ginsenosides. *Metab. Eng.* 67, 104–111. doi:10.1016/j.ymben.2021.06.002

Sibirny, A. A. (2016). Yeast peroxisomes: Structure, functions and biotechnological opportunities. *FEMS Yeast Res.* 16 (4), fow038. doi:10.1093/femsyr/fow038

Son, S. H., Kim, J. E., Moon, S. Y., Jang, I. S., Yu, B. J., and Lee, J. Y. (2022a). Metabolic recycling of storage lipids promotes squalene biosynthesis in yeast. *Biotechnol. Biofuels Bioprod.* 15 (1), 108. doi:10.1186/s13068-022-02208-9

Son, S. H., Park, G., Lim, J., Son, C. Y., Oh, S. S., and Lee, J. Y. (2022b). Chain flexibility of medicinal lipids determines their selective partitioning into lipid droplets. *Nat. Commun.* 13 (1), 3612. doi:10.1038/s41467-022-31400-6

Spanova, M., Czabany, T., Zellnig, G., Leitner, E., Hapala, I., and Daum, G. (2010). Effect of lipid particle biogenesis on the subcellular distribution of squalene in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285 (9), 6127–6133. doi:10.1074/jbc.M109.074229

Teixeira, P. G., David, F., Siewers, V., and Nielsen, J. (2018). Engineering lipid droplet assembly mechanisms for improved triacylglycerol accumulation in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 18 (6). doi:10.1093/femsyr/foy060

Tower, R. J., Fagarasanu, A., Aitchison, J. D., and Rachubinski, R. A. (2011). The peroxin Pex34p functions with the Pex11 family of peroxisomal divisional proteins to regulate the peroxisome population in yeast. *Mol. Biol. Cell* 22 (10), 1727–1738. doi:10.1091/mbc.E11-01-0084

Valm, A. M., Cohen, S., Legant, W. R., Melunis, J., Hershberg, U., Wait, E., et al. (2017). Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* 546 (7656), 162–167. doi:10.1038/nature22369

van der Klei, I. J., and Veenhuis, M. (1997). Yeast peroxisomes: Function and biogenesis of a versatile cell organelle. *Trends Microbiol.* 5 (12), 502–509. doi:10.1016/s0966-842x(97)01156-6

Verma, P., Mathur, A. K., Srivastava, A., and Mathur, A. (2012). Emerging trends in research on spatial and temporal organization of terpenoid indole alkaloid pathway in Catharanthus roseus: A literature update. *Protoplasma* 249 (2), 255–268. doi:10.1007/s00709-011-0291-4

Vizeacoumar, F. J., Torres-Guzman, J. C., Bouard, D., Aitchison, J. D., and Rachubinski, R. A. (2004). Pex30p, Pex31p, and Pex32p form a family of peroxisomal integral membrane proteins regulating peroxisome size and number in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 15 (2), 665–677. doi:10.1091/mbc.e03-09-0681

Vizeacoumar, F. J., Torres-Guzman, J. C., Tam, Y. Y., Aitchison, J. D., and Rachubinski, R. A. (2003). YHR150w and YDR479c encode peroxisomal integral membrane proteins involved in the regulation of peroxisome number, size, and distribution in *Saccharomyces cerevisiae*. *J. Cell Biol.* 161 (2), 321–332. doi:10.1083/jcb.200210130

Wei, L.-J., Cao, X., Liu, J.-J., Kwak, S., Jin, Y.-S., Wang, W., et al. (2021). Increased accumulation of squalene in engineered yarrowia lipolytica through deletion of PEX10 and URE2. *Appl. Environ. Microbiol.* 87 (17), 00481211–e100421. doi:10.1128/AEM.00481-21

Wei, L. J., Kwak, S., Liu, J. J., Lane, S., Hua, Q., Kweon, D. H., et al. (2018). Improved squalene production through increasing lipid contents in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 115 (7), 1793–1800. doi:10.1002/bit.26595

Wu, T., Ye, L., Zhao, D., Li, S., Li, Q., Zhang, B., et al. (2017). Membrane engineering - a novel strategy to enhance the production and accumulation of β-carotene in *Escherichia coli*. *Metab. Eng.* 43, 85–91. doi:10.1016/j.ymben.2017.07.001

Ye, L., Zhu, X., Wu, T., Wang, W., Zhao, D., Bi, C., et al. (2018). Optimizing the localization of astaxanthin enzymes for improved productivity. *Biotechnol. Biofuels* 11, 278. doi:10.1186/s13068-018-1270-1

Yee, D. A., DeNicola, A. B., Billingsley, J. M., Creso, J. G., Subrahmanyam, V., and Tang, Y. (2019). Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. *Metab. Eng.* 55, 76–84. doi:10.1016/j.ymben.2019.06.004

Yocum, H. C., Pham, A., and Da Silva, N. A. (2021). Successful enzyme colocalization strategies in yeast for increased synthesis of non-native products. *Front. Bioeng. Biotechnol.* 9, 606795. doi:10.3389/fbioe.2021.606795

Yofe, I., Soliman, K., Chuartzman, S. G., Morgan, B., Weill, U., Yifrach, E., et al. (2017). Pex35 is a regulator of peroxisome abundance. *J. Cell Sci.* 130 (4), 791–804. doi:10.1242/jcs.187914

Yu, Y., Rasool, A., Liu, H., Lv, B., Chang, P., Song, H., et al. (2020). Engineering *Saccharomyces cerevisiae* for high yield production of α-amyrin via synergistic remodeling of α-amyrin synthase and expanding the storage pool. *Metab. Eng.* 62, 72–83. doi:10.1016/j.ymben.2020.08.010

Yuan, J., and Ching, C. B. (2016). Mitochondrial acetyl-CoA utilization pathway for terpenoid productions. *Metab. Eng.* 38, 303–309. doi:10.1016/j.ymben.2016.07.008

Zhang, C., and Hong, K. (2020). Production of terpenoids by synthetic biology approaches. *Front. Bioeng. Biotechnol.* 8, 347. doi:10.3389/fbioe.2020.00347

Zhang, G., Wang, H., Zhang, Z., Verstrepen, K. J., Wang, Q., and Dai, Z. (2022). Metabolic engineering of yarrowia lipolytica for terpenoids production: Advances and perspectives. *Crit. Rev. Biotechnol.* 42 (4), 618–633. doi:10.1080/07388551.2021.1947183

Zhang, L., Xiao, W.-H., Wang, Y., Yao, M.-D., Jiang, G.-Z., Zeng, B.-X., et al. (2017). Chassis and key enzymes engineering for monoterpenes production. *Biotechnol. Adv.* 35 (8), 1022–1031. doi:10.1016/j.biotechadv.2017.09.002

Zhang, Y., Wang, J., Cao, X., Liu, W., Yu, H., and Ye, L. (2020). High-level production of linalool by engineered *Saccharomyces cerevisiae* harboring dual mevalonate pathways in mitochondria and cytoplasm. *Enzyme Microb. Technol.* 134, 109462. doi:10.1016/j.enzmictec.2019.109462

Zhu, K., Kong, J., Zhao, B., Rong, L., Liu, S., Lu, Z., et al. (2021). Metabolic engineering of microbes for monoterpenoid production. *Biotechnol. Adv.* 53, 107837. doi:10.1016/j.biotechadv.2021.107837

Zhu, Z.-T., Du, M.-M., Gao, B., Tao, X.-Y., Zhao, M., Ren, Y.-H., et al. (2021). Metabolic compartmentalization in yeast mitochondria: Burden and solution for squalene overproduction. *Metab. Eng.* 68, 232–245. doi:10.1016/j.ymben.2021.10.011

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 02 March 2023
DOI 10.3389/fbioe.2023.1118948

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Yu Wang,
Tianjin Institute of Industrial Biotechnology (CAS), China
Pablo Carbonell,
Universitat Politècnica de València, Spain

*CORRESPONDENCE
Quanfeng Liang,
liangquanfeng@sdu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 08 December 2022
ACCEPTED 21 February 2023
PUBLISHED 02 March 2023

CITATION
Hao R, Wang S, Jin X, Yang X, Qi Q and Liang Q (2023), Dynamic and balanced regulation of the *thrABC* operon gene for efficient synthesis of L-threonine. *Front. Bioeng. Biotechnol.* 11:1118948. doi: 10.3389/fbioe.2023.1118948

Dynamic and balanced regulation of the *thrABC* operon gene for efficient synthesis of L-threonine

Ruxin Hao†, Sumeng Wang†, Xin Jin, Xiaoya Yang, Qingsheng Qi and Quanfeng Liang*

State Key Laboratory of Microbial Technology, Shandong University, Jinan, China

L-threonine is an essential amino acid used widely in food, cosmetics, animal feed and medicine. The *thrABC* operon plays an important role in regulating the biosynthesis of L-theronine. In this work, we systematically analyzed the effects of separating *thrAB* and *thrC* in different proportions on strain growth and L-threonine production in *Escherichia coli* firstly. The results showed that higher expression of *thrC* than *thrAB* enhanced cell growth and L-threonine production; however, L-threonine production decreased when the *thrC* proportion was too high. The highest L-threonine production was achieved when the expression intensity ratio of *thrAB* to *thrC* was 3:5. Secondly, a stationary phase promoter was also used to dynamically regulate the expression of engineered *thrABC*. This strategy improved cell growth and shortened the fermentation period from 36 h to 24 h. Finally, the acetate metabolic overflow was reduced by deleting the *ptsG* gene, leading to a further increase in L-threonine production. With these efforts, the final strain $P_{2.1}$-2901Δ*ptsG* reached 40.06 g/L at 60 h fermentation, which was 96.85% higher than the initial control strain TH and the highest reported titer in shake flasks. The maximum L-threonine yield and productivity was obtained in reported fed-batch fermentation, and L-threonine production is close to the highest titer (127.30 g/L). In this work, the expression ratio of genes in the *thrABC* operon in *E. coli* was studied systematically, which provided a new approach to improve L-threonine production and its downstream products.

KEYWORDS

L-Threonine, *thrABC*, balance, dynamic, *ptsG*

Introduction

As the basic unit of proteins, amino acids play an important role in the nutrition and health maintenance of humans and animals (Lee et al., 2013). Among the 20 amino acids, L-threonine, as one of the nine essential amino acids, is used widely in food, cosmetics, animal feed and medicine. The market demand for L-threonine produced by microbial fermentation is increasing because of the growing demand for this amino acid (Wendisch, 2020). Amino acid synthesis starts at the glycolytic pathway when glucose is used as the carbon source. L-threonine biosynthesis uses the TCA cycle intermediate oxaloacetate or fumarate as a precursor. Aspartate transaminase, encoded by *aspC*, initially catalyzes oxaloacetate or aspartase, encoded by *aspA*, catalyzes fumarate to L-aspartate (Mu et al., 2021). Starting from L-aspartate, the L-threonine pathway consists of five enzymes that catalyze the reaction: aspartokinase encoded by *thrA*, *metL* or *lysC*; semialdehyde dehydrogenase encoded by *asd*; homoserine dehydrogenase encoded by *thrA*;

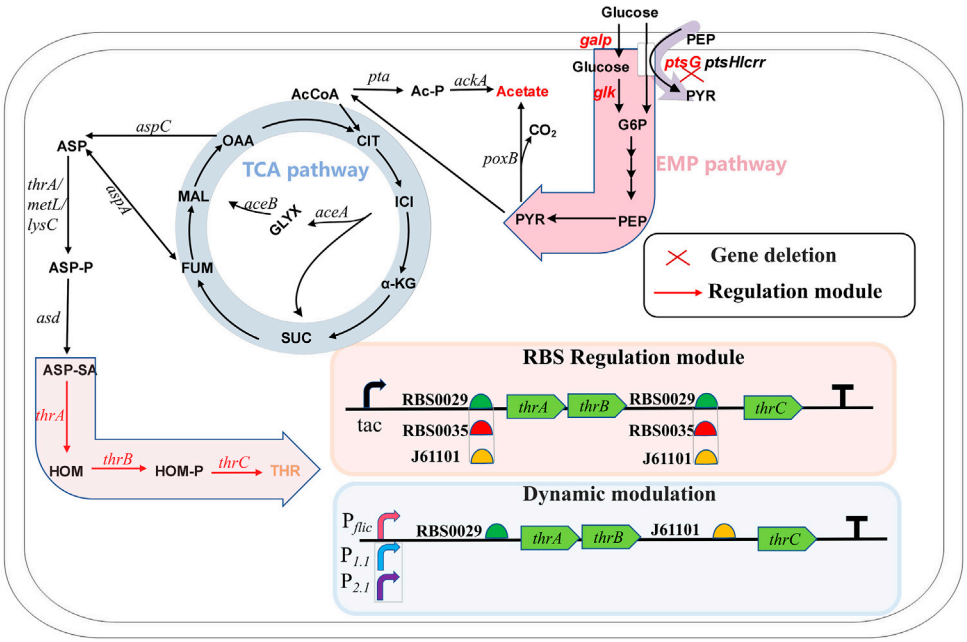


FIGURE 1
The metabolic pathway of L-threonine production in *Escherichia coli* and the modification strategy used in this study. G6P, glucose-6-phosphate; PEP, phosphoenol pyruvate; PYR, pyruvate; AcCoA, acetyl-coenzyme A; Ac-P, phosphorylated acetyl coenzyme A; CIT, citrate; ICI, isocitrate; GLYX, glyoxylate; α-KG, α-ketoglutarate; SUC, succinate; FUM, fumarate; MAL, malate; OAA, oxaloacetate; ASP, aspartate; ASP-P, aspartate phosphate; ASP-SA, aspartate semialdehyde; HOM, L-homoserine; HOM-P, homoserine phosphate; THR, L-threonine.

homoserine kinase encoded by *thrB*; and threonine synthase encoded by *thrC* (Wang et al., 2022a) (Figure 1). Many strategies for synthesizing L-threonine have been explored, including overexpression of key genes, elimination of competing pathways and feedback inhibition, regulation of capacity and increased extracellular transport of L-threonine (Lee et al., 2007; Dong et al., 2011; Mu et al., 2021).

In *Escherichia coli*, *thrA*, *thrB* and *thrC* are arranged adjacently on the chromosome to form the *thrABC* operon. The *thrABC* operon controls several key enzymes, from L-Aspartate-4-semialdehyde to L-threonine synthesis. Research on *thrABC* expression and regulation is important for L-threonine synthesis and the downstream products, L-isoleucine and L-glycine. Overexpression of the *thrABC* operon in a bacterial strain increased L-threonine production significantly (Lee et al., 2003; Zhao et al., 2018), and an L-isoleucine-producing strain was constructed by overexpressing *thrABC* (Park et al., 2012). Threonine synthase encoded by *thrC* catalyzes the production of L-threonine from homoserine phosphate, which is the last step in the synthesis of L-threonine. Lee et al. reported that this reaction step is slower than reactions catalyzed upstream, and they studied how assembling homoserine dehydrogenase (HDH), homoserine kinase (HK) and different amounts of threonine synthase (TS) on the DNA scaffold affected the production rate of L-threonine (Lee et al., 2013). However, no study has examined the effect of reducing TS expression below that of HDH and HK. Recent studies have shown that regulating the ratio of different genes on the same operon is crucial for the efficient synthesis of products (Zelcbuch et al., 2013). Therefore, systematically studying the effect of varying the expression ratio of the *thrABC* operon genes may improve L-threonine production.

Dynamic regulation can prevent metabolic burden on cells caused by constitutive gene expression in L-threonine production. We previously designed a positive feedback strategy to dynamically regulate the expression of threonine transporters, which alleviated the adverse effects of overexpressed transporters on cells and excreted intracellular metabolites (Wang et al., 2022a). In addition, promoters of *cysH*, *cysJ* and *cysD* that can be activated by L-threonine control the expression of *aspC* gene, which did not affect cell growth and increased the production of L-threonine (Zhao et al., 2020).

In this study, we developed a new strategy for dynamic and balanced regulation of the *thrABC* operon for efficient synthesis of L-threonine (Figure 1). We first explored the optimal expression levels of *thrAB* and *thrC* to improve L-threonine production. We then used dynamic regulation to control the expression of *thrABC*. Finally, we deleted the PTS system in an effort to reduce metabolic overflow and further increase L-threonine production.

Materials and methods

Construction of strains and plasmids

The strains and plasmids used in this study are listed in Supplementary Tables S1, S2. *E. coli* DH5α was used for molecular cloning and manipulation of plasmids. All primers used are listed in Supplementary Table S3. All plasmids were

constructed using the Gibson assembly method (Gibson et al., 2009). *E. coli* K-12 MG1655 was used for RBS strength characterization. RBS (B0029, B0030, B0031, B0032, B0033, B0034, B0035, B0064 and J61101) of different strengths were analyzed by the ratio of RFP to OD_{600}. Nine plasmids were assembled. The expression of *thrAB* and *thrC* was regulated by using the primers listed in Supplementary Table S3, which replaced various RBS to construct L-threonine production plasmids (PA-29*thrAB*-29*thrC*, PA-29*thrAB*-01*thrC*, PA-29*thrAB*-35*thrC*, PA-01*thrAB*-29*thrC*, PA-01*thrAB*-01*thrC*, PA-01*thrAB*-35*thrC*, PA-35*thrAB*-29*thrC*, PA-35*thrAB*-01*thrC* and PA-35*thrAB*-35*thrC*). The TH strain producing L-threonine was provided by the Fufeng Group (Qingdao, China) and derived from *E. coli* K-12 MG1655. Three genes *thrA*, *thrB* and *thrC* increase from one copy number to four copies in the genome. Comparison of partial genes of strains TH and MG1655 are listed in Supplementary Table S4. The expression level of *thrAB* and *thrC* in L-threonine production was balanced by transforming the above plasmids into strain TH to generate strains 2929, 2935, 2901, 0129, 0101, 0135, 3529, 3535 and 3501. P_{flic}, $P_{1.1}$ and $P_{2.1}$ promoter sequences were derived from reports (Jaishankar and Srivastava, 2020; Zhang et al., 2021). The *ptsG* gene deletion experiment was carried out by homologous recombination (Datsenko and Wanner, 2000). All fragments were amplified with Phanta HS ultra-fidelity DNA polymerase (Vazyme Biotech, Nanjing, China).

RBS characterization

RBS characterization was achieved by monitoring the RFP fluorescence and cell density in real-time using a multiple detection microplate analyzer (SynergyHT, BioTek, Winooski, VT, United States). The details are as follows. The seeds of RBS characterization strains were prepared by transferring a single colony to a 12-well microassay plate containing 2 mL LB medium with 34 μg/mL chloramphenicol. The cells were grown at 37°C for 12 h. Next, 2% (v/v) of the seeds were inoculated into 0.2 mL LB medium with 34 μg/mL chloramphenicol in a 96-well microassay plate to detect red fluorescence. The 96-well plate was incubated at 37°C with oscillation. During RBS characterization, strain growth was measured at 600 nm. The red fluorescence was detected by excitation at 590 nm and emission at 645 nm.

Medium and L-threonine fermentation

LB medium was used for plasmid construction and RBS strength characterization. L-threonine fermentation medium consists of 15 g/L $(NH_4)_2SO_4$, 2 g/L KH_2PO_4, 1 g/L $MgSO_4{\cdot}7H_2O$, 2 g/L yeast extract and 0.02 g/L FeSO4 (Kruse et al., 2002). Glucose (40 g/L) was added as the initial carbon source and 20 g/L $CaCO_3$ was used to adjust the pH during fermentation. For shake flask fermentation, a single colony was incubated in fresh LB medium at 37°C for 12 h. The precultured seeds were then transferred with 1% (v/v) inoculation to a 300 mL shake flask that contained 20 mL fermentation medium. Fermentation was carried out at 220 rpm and 37°C. Cultures were supplemented with 40 g/L glucose when the glucose level was lower than 15 g/L.

The fed-batch culture was carried out in a 5-L bioreactor containing 4L medium (20 g/L $(NH_4)_2SO_4$, 3 g/L yeast extract, 2 g/L KH_2PO_4, 2 g/L $MgSO_4{\cdot}7H_2O$, 5 mg/L $FeSO_4{\cdot}7H_2O$, 5 mg/L $MnSO_4{\cdot}4H_2O$, 0.5 g/L betaine). Temperature was maintained at 37°C, the aeration rate at 1.5 vvm, pH was maintained automatically at 7.0 with NH_4OH, and the dissolved oxygen value was maintained below 30%. 20 g/L initial amount of sterilized glucose was added in the working culture, and glucose concentration was controlled by continuous feeding.

Analytical methods

One milliliter of the culture was mixed vigorously, and 0.1 mL was transferred to a 1.5 mL centrifuge tube. One millimolar HCl (0.9 mL) was added to this culture, and the sample was mixed to remove residual $CaCO_3$. Subsequently, the OD_{600} was measured using a spectrophotometer (Shimadzu, Kyoto, Japan). For glucose and acetate assays, the cultures were centrifuged at 12,000 rpm for 2 min to collect the supernatant. The collected supernatant was filtered through a 0.22 μm water membrane for analysis. A refractive index detector (RID-10A; Shimadzu, Kyoto, Japan) and an AminexHPX-87H ion exclusion column (Bio-Rad Laboratories, Hercules, CA, United States) were used with 5 mM H_2SO_4 as the mobile phase and a flow rate of 0.6 mL/min.

For the detection of L-threonine, the collected supernatant was deproteinized with 5% trichloroacetic acid. Subsequently, the pretreated supernatant was derivatized with triethylamine and phenyl isothiocyanate, followed by extraction with *n*-hexane (Wang et al., 2022a). Briefly, 0.2 mL of sample and standard L-threonine were pretreated with a mixture of triethylamine-acetonitrile (1.4 mL of triethylamine mixed with 8.6 mL of acetonitrile). Next, we added phenylisothiocyanate-acetonitrile (25 μL of phenylisothiocyanate mixed with 2 mL of acetonitrile) to pretreat samples and the L-threonine standard for 1 h at room temperature. *n*-Hexane (0.4 mL) was added, and the sample was shaken vigorously. The lower layer (0.2 mL) was collected and diluted with 0.8 mL deionized water. The solution was filtered with a 0.22 μm organic membrane, and samples were detected using an HPLC equipped with a diode array detector (SPD-M20A; Shimadzu, Kyoto, Japan) and a VenusilAA (4.6 × 250 mm, 5 μm, AgelaTechanology) column at 40°C. The mobile phase consisted of (A) 15.2 g sodium acetate dissolved in 1850 mL of ultrapure water and mixed with 140 mL of acetonitrile and (B) 80% (v/v) acetonitrile and 20% (v/v) ultrapure water. The flow rate was 1 mL/min. The L-threonine concentration was quantified using the corresponding standard curve and peak area.

Results and discussion

Balancing the expression of *thrAB* and *thrC* to promote cell growth and L-threonine production

In the biosynthesis of L-threonine in *E. coli*, genes *thrA*, *thrB* and *thrC* encoding the last three key enzymes of this biosynthesis are located on the *thrABC* operon. ThrC catalyzes the final step of

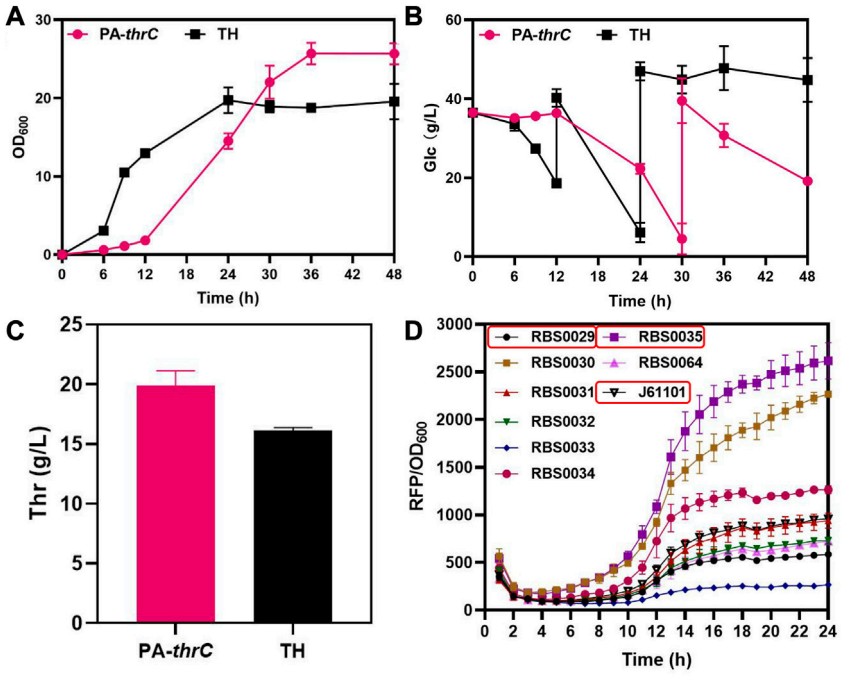


FIGURE 2
Comparison of the fermentation results of overexpressed *thrC* by medium copy PACYC-Duet plasmid and control TH and characterization results of different RBS intensities. The PA-ThrC strain overexpresses *thrC* on the PACYC-Duet plasmid. All results were derived from three (n = 3) independent repeats. **(A)** Comparison of cell growth (OD_{600}) between overexpressing *thrC* and control strains. **(B)** Glucose consumption by the overexpressing *thrC* and control strains. **(C)** Comparison of L-threonine titer between the overexpressing *thrC* strain and the control strain after fermentation for 48 h. **(D)** The fluorescence intensity of RFP was monitored in real-time by a multi-detection microplate reader to characterize and screen RBS with different intensities. The RBS intensity was calculated using the ratio of RFP to OD_{600}.

L-threonine synthesis, and this reaction is slow (Lee et al., 2013). As the rate-limiting step of L-threonine production, it is necessary to enhanced the expression of *thrC*. We used the L-threonine-producing strain TH available from FuFeng Group as the initial strain, which was derived from MG1655 (Supplementary Table S2). Initially, we enhanced the expression of *thrC* on the low copy PCL-1920 plasmid and transformed the plasmid into the initial strain TH to obtain the pcl-*thrC* strain. However, the production of L-threonine and cell growth of the strain was not significantly improved compared with the original strain TH (Supplementary Figure S1A,B). We analyzed that the reason is the weak degree of enhanced expression. Next, we enhanced the expression of *thrC* on the medium copy PACYC-Duet plasmid and transformed the plasmid into the initial strain TH to obtain the PA-*thrC* strain. Although the production of L-threonine increased, cell growth and the glucose consumption rate were seriously impeded during early cultivation (Figures 2A–C). In addition, *thrA*, *thrB*, *thrC* are located on the *thrABC* operon. So the overexpressed *thrC* might disrupt the balance between *thrAB* and *thrC*. Therefore, we analyzed the influence of various ratios of *thrAB* and *thrC* on L-threonine synthesis and cell growth. Ribosome binding sites (RBS) define the translation efficiency of genes and are typically used in gene regulation studies (Hur et al., 2020; Wang et al., 2022b). We selected three RBS with different intensities by fluorescence characterization, RBS0035 (2615), J61101 (958) and RBS0029 (585), to regulate *thrAB* and *thrC* expression levels, with an intensity ratio between them of 13:5:3 (Figure 2D). We constructed nine strains containing various combinations and levels of *thrAB* and *thrC* expression.

The results showed that different combinations of *thrAB* and *thrC* expression levels had different effects on cell growth. High expression of *thrC* promoted growth and L-threonine accumulation simultaneously; however, there were also ratios between *thrAB* and *thrC* when the level of L-threonine decreased because the proportion of *thrC* was too high. Among the nine strains constructed, the 2901 strain (*thrAB*:*thrC* = 3:5) yielded the highest L-threonine titer. The 2935 strain (3:13) grew slowly because of the large difference in the expression levels of *thrAB* and *thrC*. Optimal growth was observed for the 0135 strain (5:13) (Figure 3A; Figure 4). Accumulation of L-threonine was poor when *thrC* was expressed weakly, and different ratios had minimal effect on L-threonine production (Figure 3B, Figure 4). Equal expression levels of *thrAB* and *thrC* in strain 3535 yielded a relatively good L-threonine titer, but production was lower than that of strain 2901 (Figure 3C, Figure 4). The results indicated that high expression of *thrC* in an appropriate *thrAB*:*thrC* ratio yielded optimal cell growth and L-threonine production.

In summary, the results showed that when the expression of *thrC* was higher than *thrAB*, it was beneficial to the growth and production of L-threonine, and the level of L-threonine decreased when the expression of *thrC* was much higher than *thrAB* in particular strains. Using a *thrAB*:*thrC* expression ratio of 3:5 yielded the highest production of L-threonine. Production of L-threonine by proportional expression of *thrAB* and *thrC* was

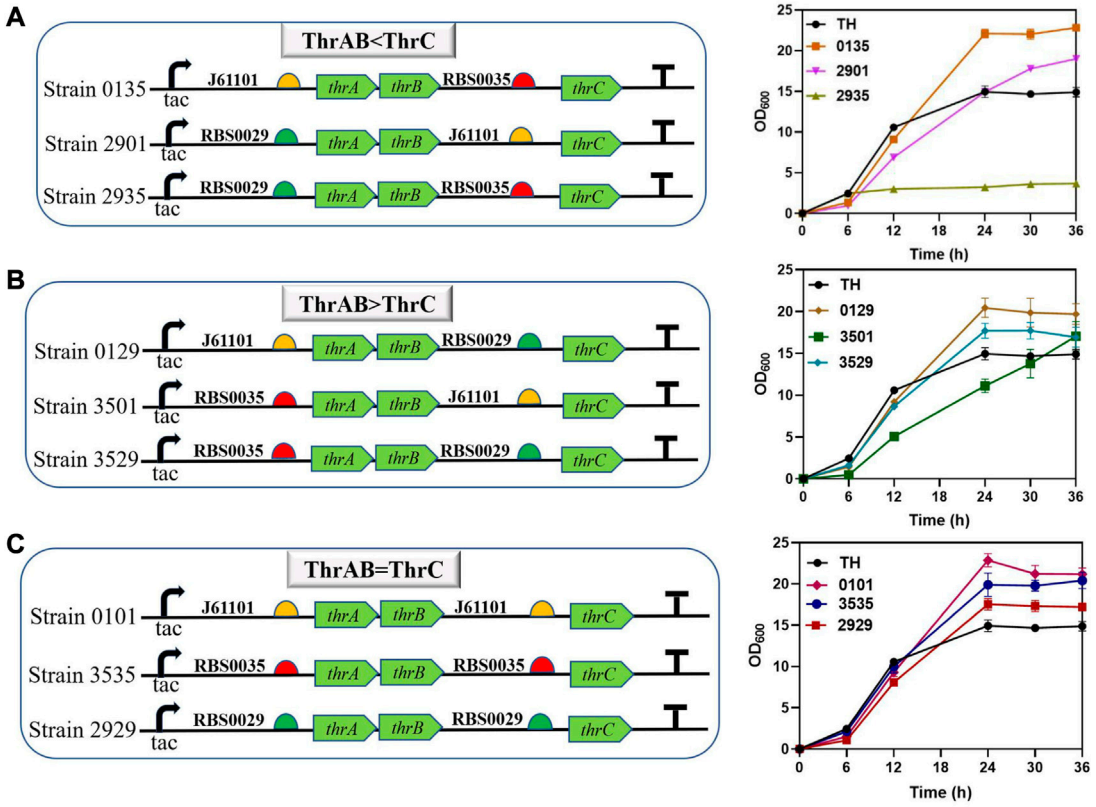


FIGURE 3
The effect of different expression intensities of *thrAB* and *thrC* on strain TH growth. According to the RBS strength, the adjusted expression of *thrAB* and *thrC* was classified into three groups: ThrAB < ThrC, ThrAB > ThrC and ThrAB = ThrC. All results were calculated from three (n = 3) independent repeats. **(A)** Strains with *thrAB* expression weaker than *thrC* were constructed using different combinations of RBS, and fermentation results OD_{600} are shown. **(B)** Strains with *thrAB* expression stronger than *thrC* were constructed by using different combinations of RBS, and fermentation results OD_{600} are shown. **(C)** Strains with equal expression strengths of *thrAB* and *thrC* were constructed by using different combinations of RBS, and fermentation results OD_{600} are shown.

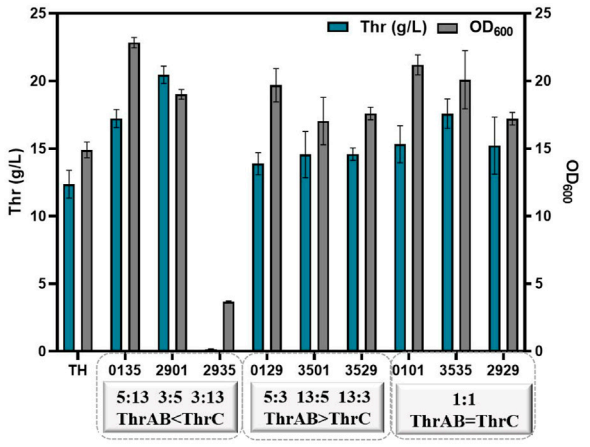


FIGURE 4
The effect of different expression intensities of *thrAB* and *thrC* on L-threonine production by the TH strain. All results were derived from three (n = 3) independent repeats.

found to be a more effective approach to produce L-threonine than direct overexpression of *thrABC* (Zhao et al., 2018).

Dynamic regulation of the engineered *thrABC* to counter metabolic burden and increase L-threonine productivity

As described above, L-threonine production was promoted by carefully regulating the expression levels of *thrAB* and *thrC* and optimizing the *thrAB*:*thrC* ratio; however, constitutive overexpression of the *thrABC* operon inhibited cell growth (Figures 3A–C). We adopted a dynamic regulation strategy to ensure that the metabolic burden caused by the premature introduction of the *thrABC* operon was alleviated. Instead of using IPTG and other chemical inducers, a growth-related promoter was used to control *thrABC* expression after obtaining a high cell density. Several stationary phase promoters with different strengths have been obtained previously by random mutagenesis of a wild-type stationary phase promoter (Shimada et al., 2004). The

10.3389/fbioe.2023.1118948

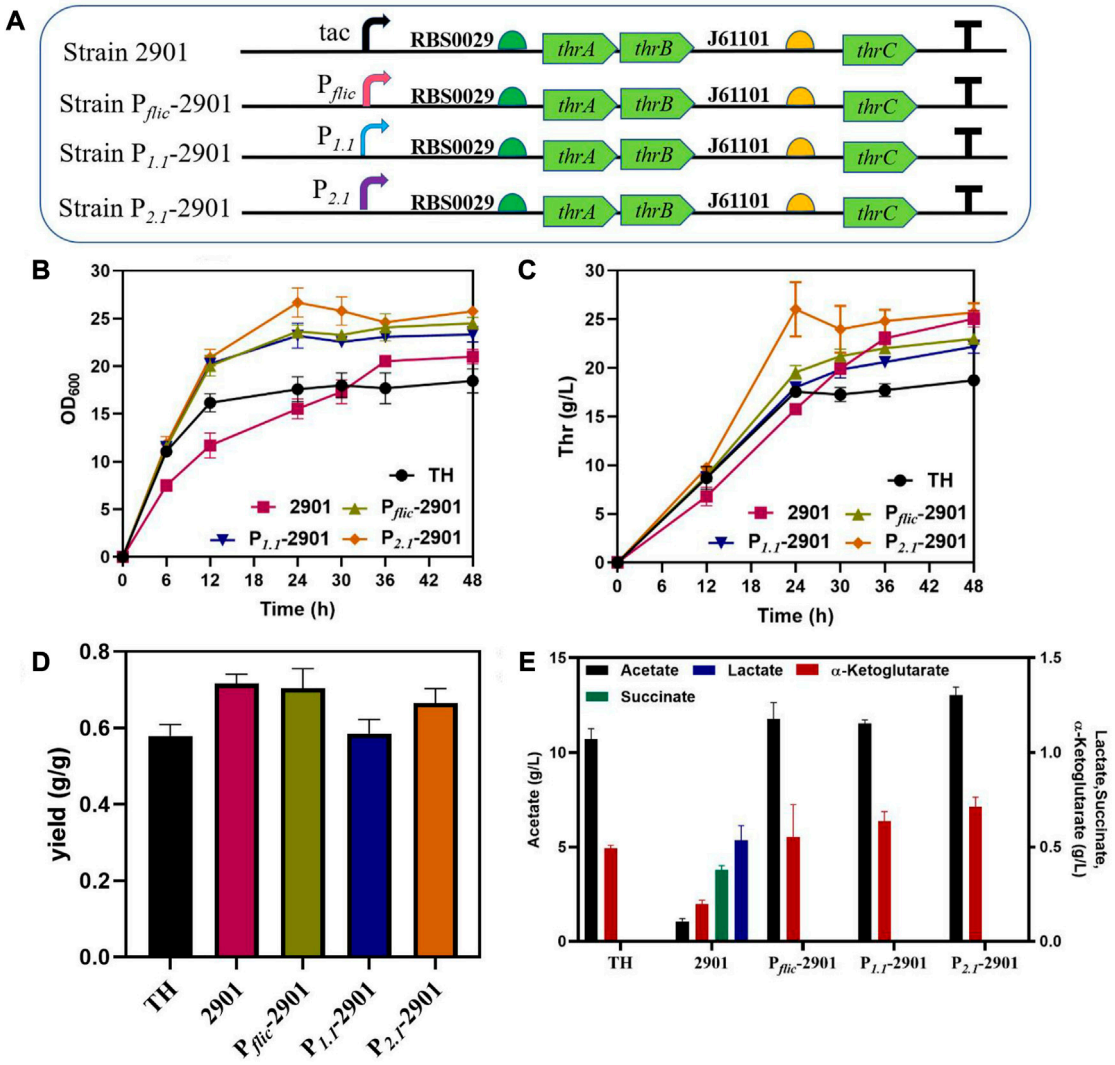


FIGURE 5
The effect of dynamic control of the *thrABC* operon with a stationary phase promoter on the growth and L-threonine production in *Escherichia coli* TH. All results were calculated from three (n = 3) independent repeats. **(A)** The strain was constructed by replacing the tac promoter with different promoters. **(B)** The OD_{600} of fermenting strains is controlled by a stationary phase promoter. **(C)** Samples were taken at 12, 24, 30, 36 and 48 h during fermentation to measure L-threonine production. **(D)** The yield of the final product was calculated by the ratio of glucose consumption to L-threonine titer. **(E)** The content of the by-product acetate and other metabolites at the end of fermentation for 48 h.

strength of the $P_{1.1}$ and $P_{2.1}$ promoters obtained was equivalent to *E. coli* promoters (Jaishankar and Srivastava, 2020). The promoters of *fliA*, *fliC* and *flgC* involved in flagella construction were also identified as stationary phase promoters. The transcriptional intensity of these promoters is active during the lag and early exponential phases but inhibited throughout the late exponential and stationary phases (Zhang et al., 2019).

The effect of dynamic regulation of *thrABC* on strain growth and production was verified by using three stationary phase promoters, P_{fliC}, $P_{1.1}$ and $P_{2.1}$, for *thrABC* regulation in the PA-29*thrAB*-01*thrC* plasmid (Figure 5A). The results showed that the strains controlled by the three stationary phase promoters displayed improved growth in the early stage, and L-threonine production increased significantly compared with the control group. Among these strains, growth was strongest for strain $P_{2.1}$-2901, which was controlled by $P_{2.1}$, and L-threonine production was similar to that of the strain controlled by the tac promoter (Figures 5B,C). This observation is because the $P_{2.1}$ promoter is the strongest among the three tested (Jaishankar and Srivastava, 2020). We measured L-threonine production in each period and found that L-threonine accumulation in the $P_{2.1}$-2901 strain was the fastest, with the L-threonine titer and productivity reaching 26.02 g/L and 1.08 g/L/h at 24 h, respectively. L-threonine production by strain $P_{2.1}$-2901 was 44.56% and 65.00% higher than that of the control strain TH and strain 2901 controlled by the tac promoter, respectively (Figure 5C). This result arises because the replaced stationary phase promoter is more active during the early exponential phase, which promotes the overexpression of the *thrABC* operon to accelerate the accumulation of L-threonine. This phenomenon suggests that this strategy can enhance product titer and shorten the fermentation

10.3389/fbioe.2023.1118948

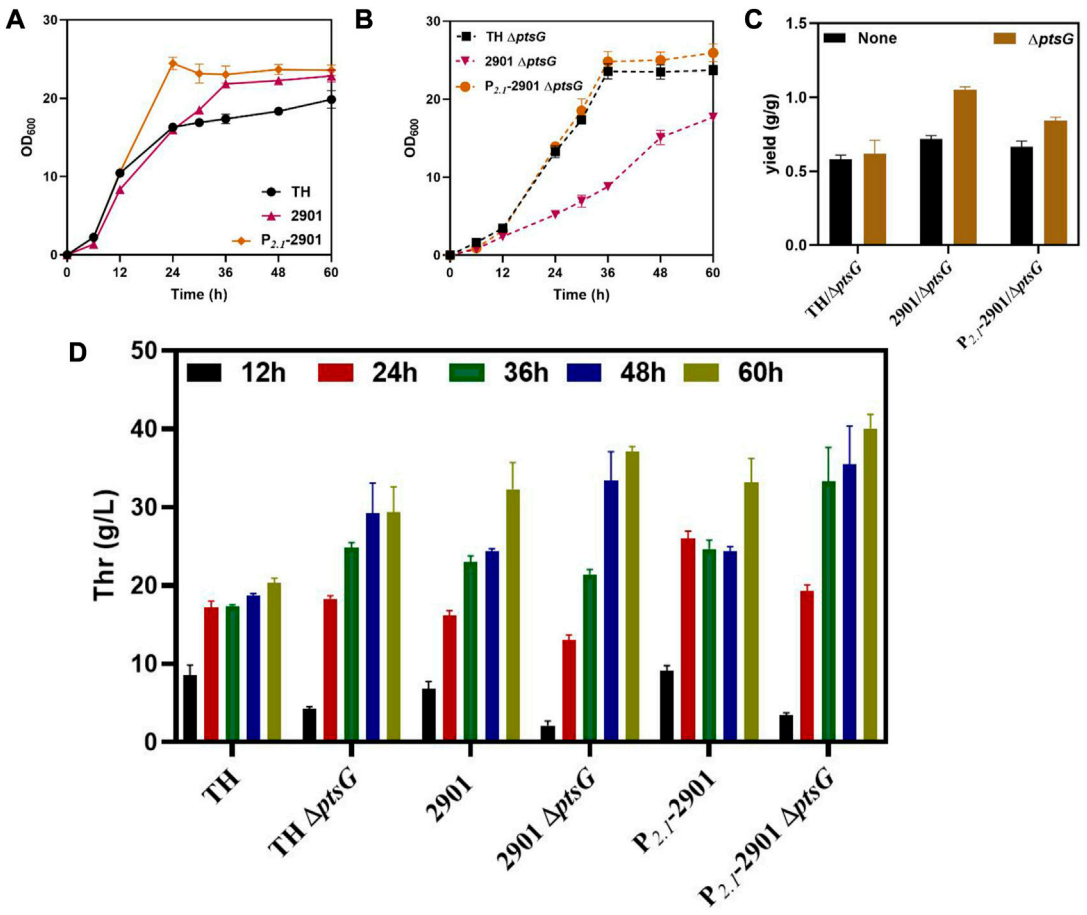


FIGURE 6
The effect of deleting *ptsG* on the growth and L-threonine production of strain TH. All results were determined from three (n = 3) independent repeats. **(A)** Growth of strains before *ptsG* gene deletion. **(B)** Growth of strains after *ptsG* gene deletion. **(C)** The yield of the final product was calculated by the glucose consumption and L-threonine titer. **(D)** During fermentation, samples taken at 12, 24, 36, 48 and 60 h were used to measure L-threonine production.

period. The overexpression of the *thrABC* gene controlled by the stationary phase promoter can improve L-threonine yield (Figure 5D).

Although we increased L-threonine productivity by using a stationary phase promoter, the strains with better growth in the early stage also produced a large amount of acetate (Figure 5E). Studies have suggested that the rapid use of glucose during the early stage of cell growth will lead to an acetate overflow (Eiteman and Altman, 2006). Thus, we next tried to reduce the acetate overflow.

Decreasing glucose transport to reduce acetate overflow and improve L-threonine production

The phosphotransferase system (PTS) is the major glucose transport system in *E. coli*. Deletion of PTS has been reported to reduce acetate overflow and increase the production of recombinant proteins and biochemicals (Wang et al., 2006; Kang et al., 2009). We deleted the *ptsG* gene in three threonine-producing strains, TH, 2901 and $P_{2.1}$-2901, to examine the effect of PTS activity on L-threonine production.

The growth of the three strains was retarded after deleting *ptsG*, which was caused by the decrease in the glucose uptake rate by the cells. The growth of the 2901Δ*ptsG* strain was inhibited significantly because of the combined effects of *thrABC* overexpression and *ptsG* deletion. Optimal growth was observed for strain $P_{2.1}$-2901Δ*ptsG*, reaching a stable growth state after 36 h, whereas the 2901Δ*ptsG* strain showed slower growth up to 60 h (Figures 6A,B). This observation also reflects the advantages of dynamic regulation. Using the $P_{2.1}$-2901Δ*ptsG* strain can shorten the fermentation period and increase L-threonine production. L-threonine production by the three strains was improved when compared with the corresponding strains without the *ptsG* deletion, with an increase in L-threonine yield (Figure 6C). At 60 h, L-threonine production by strain $P_{2.1}$-2901Δ*ptsG* reached 40.06 g/L, which was 36.30% and 7.95% higher than that of THΔ*ptsG* and 2901Δ*ptsG* strains, respectively (Figure 6D). As expected, acetate was not generated during fermentation.

To further verify the properties of $P_{2.1}$-2901Δ*ptsG* for L-threonine production, fed-batch fermentation was performed in 5-L bioreactor,

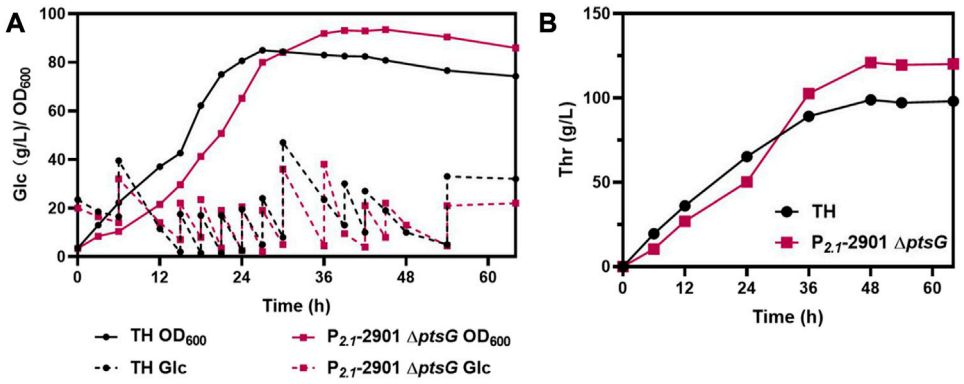


FIGURE 7
The fed-batch process with $P_{2.1}$-2901Δ*ptsG* and TH in a 5-L bioreactor. Biomass, glucose consumption, and L-threonine titer were monitored in real time. **(A)** Biomass and glucose consumption of strains. **(B)** L-threonine titer of strains.

TABLE 1 L-threonine production using different strategies.

Species	Shake flasks fermentation			Fed-batch fermentation			Source
	Titer (g/L)	Yield (g/g)	Productivity (g/L/h)	Titer (g/L)	Yield (g/g)	Productivity (g/L/h)	
E. coli (TWF083)	29.73	-	-	116.62	0.49	2.43	Zhao et al. (2020)
E. coli (TWF044)	28.49	0.72	-	103.89	0.45	-	Yang et al. (2019)
E. coli (TSW009)	26.00	0.65	0.54	-	-	-	Wang et al. (2022b)
E. coli (TWF113 / pFT24rpa 1)	25.85	-	-	-	-	-	Fang et al. (2020)
E. coli (JLTHR)	-	-	-	127.30	0.58	-	Su et al. (2018)
E. coli (WMZ016/pFW01-thrA*BC-rhtC)	17.98	0.35	-	-	-	-	Zhu et al. (2019)
Halomonas bluephagenesis (TDHR3-42-p226)	7.50	-	-	33.00		1.40	Du et al. (2020)
Cotynebacterium glutamicum	12.80	-	-		-	-	Wei et al. (2018)
E. coli ($P_{2.1}$-2901Δ*ptsG*)	40.06	0.84	0.67	121.05	0.60	2.52	This Study

using TH as a control. The cell density, glucose consumption levels (Figure 7A) and L-threonine titer (Figure 7B) are depicted. Following 48 h of cultivation, $P_{2.1}$-2901Δ*ptsG* could produce 121.05 g/L L-threonine, leading to a yield of 0.60 g/g glucose and productivity of 2.52 g/L/h. TH could produce 98.88 g/L L-threonine in 48 h, with a yield and productivity of 0.49 g/g glucose and 2.06 g/L/h, respectively. After 48 h of incubation, unfavorable growth conditions and irreparable cellular damage lead to a death phase and long-term stationary phase (Finkel, 2006), resulting in the lose viability of L-threonine production and glucose consumption. In contrast to the results of other reports, $P_{2.1}$-2901Δ*ptsG* achieved the highest productivity and yield. Therefore, with respect to the three major parameters in industrial production, $P_{2.1}$-2901Δ*ptsG* possesses substantial L-threonine production ability.

The reaction catalyzed by ThrC is the rate-limiting step of L-threonine production and overexpression of *thrC* inhibits cell growth. Here, we showed that balanced regulation of *thrAB* and *thrC* expression levels promotes *E. coli* growth and L-threonine production. L-threonine production was further improved by dynamically regulating the overexpression of *thrABC*. Finally, the deletion of *ptsG* reduced acetate production and further increased the titer and yield of L-threonine. With the above approaches, we obtained strain $P_{2.1}$-2901Δ*ptsG*, with the titer of L-threonine by fermentation in a shake flask for 60 h reaching 40.06 g/L, which is 96.85% higher than that of the control strain TH (20.35 g/L) and the highest titer reported in a shake flask (Table 1). The yield reached 0.84 g/g, which was 44.82% higher than the control strain TH (0.58 g/g). $P_{2.1}$-2901Δ*ptsG* could produce 121.05 g/L L-threonine in 48 h, with a yield and productivity of 0.60 g/g glucose and 2.52 g/L/h by fed-batch fermentation, respectively. The maximum L-threonine yield and productivity was obtained in reported fed-batch fermentation, and L-threonine titer is close to the maximum (127.30 g/L) (Table 1). This study provided a new concept for improving L-threonine production and downstream products.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

RH and SW: Investigation, conceptualization, and writing original draft. XJ and XY: Data curation and formal analysis. QQ and QL: Conceptualization, funding acquisition, resource management, project administration, supervision, and writing: review and editing.

Funding

This research was funded by the National Key Research and Development Program of China (2022YFC3401300) and National Natural Science Foundation of China (31971336, 31770095).

Acknowledgments

Strain TH was kindly provided by the Fufeng Group. We thank Chengjia Zhang and Nannan Dong from the Analysis and Testing Center of the State Key Laboratory for Microbial Technology (Shandong University) for assistance in the fermentation experiment. We thank Liwen Bianji (Edanz) for editing the English text of a draft of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1118948/full#supplementary-material

References

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97 (12), 6640–6645. doi:10.1073/pnas.120163297

Dong, X., Quinn, P. J., and Wang, X. (2011). Metabolic engineering of *Escherichia coli* and Corynebacterium glutamicum for the production of L-threonine. *Biotechnol. Adv.* 29 (1), 11–23. doi:10.1016/j.biotechadv.2010.07.009

Du, H., Zhao, Y., Wu, F., Ouyang, P., Chen, J., Jiang, X., et al. (2020). Engineering Halomonas bluephagenesis for L-Threonine production. *Metab. Eng.* 60, 119–127. doi:10.1016/j.ymben.2020.04.004

Eiteman, M. A., and Altman, E. (2006). Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol.* 24 (11), 530–536. doi:10.1016/j.tibtech.2006.09.001

Fang, Y., Wang, J., Ma, W., Yang, J., Zhang, H., Zhao, L., et al. (2020). Rebalancing microbial carbon distribution for L-threonine maximization using a thermal switch system. *Metab. Eng.* 61, 33–46. doi:10.1016/j.ymben.2020.01.009

Finkel, S. E. (2006). Long-term survival during stationary phase: Evolution and the GASP phenotype. *Nat. Rev. Microbiol.* 4 (2), 113–120. doi:10.1038/nrmicro1340

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6 (5), 343–345. doi:10.1038/nmeth.1318

Hur, D. H., Choi, W. S., Kim, T. Y., Lee, S. Y., Park, J. H., and Jeong, K. J. (2020). Enhanced production of bacterial cellulose in komagataeibacter xylinus via tuning of biosynthesis genes with synthetic RBS. *J. Microbiol. Biotechnol.* 30 (9), 1430–1435. doi:10.4014/jmb.2006.06026

Jaishankar, J., and Srivastava, P. (2020). Strong synthetic stationary phase promoter-based gene expression system for *Escherichia coli*. *Plasmid* 109, 102491. doi:10.1016/j.plasmid.2020.102491

Kang, Z., Geng, Y., Xia, Y., Kang, J., and Qi, Q. (2009). Engineering *Escherichia coli* for an efficient aerobic fermentation platform. *J. Biotechnol.* 144 (1), 58–63. doi:10.1016/j.jbiotec.2009.06.021

Kruse, D., Kramer, R., Eggeling, L., Rieping, M., Pfefferle, W., Tchieu, J. H., et al. (2002). Influence of threonine exporters on threonine production in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 59 (2-3), 205–210. doi:10.1007/s00253-002-0987-7

Lee, J. H., Jung, S. C., Bui Le, M., Kang, K. H., Song, J. J., and Kim, S. C. (2013). Improved production of L-threonine in *Escherichia coli* by use of a DNA scaffold system. *Appl. Environ. Microbiol.* 79 (3), 774–782. doi:10.1128/aem.02578-12

Lee, J. H., Lee, D. E., Lee, B. U., and Kim, H. S. (2003). Global analyses of transcriptomes and proteomes of a parent strain and an L-threonine-overproducing mutant strain. *J. Bacteriol.* 185 (18), 5442–5451. doi:10.1128/jb.185.18.5442-5451.2003

Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U., and Lee, S. Y. (2007). Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* 3, 149. doi:10.1038/msb4100196

Mu, Q., Zhang, S., Mao, X., Tao, Y., and Yu, B. (2021). Highly efficient production of L-homoserine in *Escherichia coli* by engineering a redox balance route. *Metab. Eng.* 67, 321–329. doi:10.1016/j.ymben.2021.07.011

Park, J. H., Oh, J. E., Lee, K. H., Kim, J. Y., and Lee, S. Y. (2012). Rational design of *Escherichia coli* for L-isoleucine production. *ACS Synth. Biol.* 1 (11), 532–540. doi:10.1021/sb300071a

Shimada, T., Makinoshima, H., Ogawa, Y., Miki, T., Maeda, M., and Ishihama, A. (2004). Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector. *J. Bacteriol.* 186 (21), 7112–7122. doi:10.1128/jb.186.21.7112-7122.2004

Su, Y., Guo, Q. Q., Wang, S., Zhang, X., and Wang, J. (2018). Effects of betaine supplementation on L-threonine fed-batch fermentation by *Escherichia coli*. *Bioprocess Biosyst. Eng.* 41 (10), 1509–1518. doi:10.1007/s00449-018-1978-0

Wang, Q., Wu, C., Chen, T., Chen, X., and Zhao, X. (2006). Expression of galactose permease and pyruvate carboxylase in *Escherichia coli* ptsG mutant increases the growth rate and succinate yield under anaerobic conditions. *Biotechnol. Lett.* 28 (2), 203–293. doi:10.1007/s10529-006-6907-7

Wang, S., Hao, R., Jin, X., Li, X., Qi, Q., and Liang, Q. (2022a). Dynamic regulation of transporter expression to increase L-threonine production using L-threonine biosensors. *Fermentation* 8, 250. doi:10.3390/fermentation8060250

Wang, S., Jin, X., Jiang, W., Wang, Q., Qi, Q., and Liang, Q. (2022b). The expression modulation of the key enzyme acc for highly efficient 3-hydroxypropionic acid production. *Front. Microbiol.* 13, 902848. doi:10.3389/fmicb.2022.902848

Wei, L., Xu, N., Wang, Y., Zhou, W., Han, G., Ma, Y., et al. (2018). Promoter library-based module combination (PLMC) technology for optimization of threonine

biosynthesis in Corynebacterium glutamicum. *Appl. Microbiol. Biotechnol.* 102 (9), 4117–4130. doi:10.1007/s00253-018-8911-y

Wendisch, V. F. (2020). Metabolic engineering advances and prospects for amino acid production. *Metab. Eng.* 58, 17–34. doi:10.1016/j.ymben.2019.03.008

Yang, J., Fang, Y., Wang, J., Wang, C., Zhao, L., and Wang, X. (2019). Deletion of regulator-encoding genes fadR, fabR and iclR to increase L-threonine production in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 103 (11), 4549–4564. doi:10.1007/s00253-019-09818-8

Zelcbuch, L., Antonovsky, N., Bar-Even, A., Levin-Karp, A., Barenholz, U., Dayagi, M., et al. (2013). Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* 41 (9), e98. doi:10.1093/nar/gkt151

Zhang, J., Weng, H., Zhou, Z., Du, G., and Kang, Z. (2019). Engineering of multiple modular pathways for high-yield production of 5-aminolevulinic acid in *Escherichia coli*. *Bioresour. Technol.* 274, 353–360. doi:10.1016/j.biortech.2018.12.004

Zhang, Y., Wei, M., Zhao, G., Zhang, W., Li, Y., Lin, B., et al. (2021). High-level production of l-homoserine using a non-induced, non-auxotrophic *Escherichia coli* chassis through metabolic engineering. *Bioresour. Technol.* 327, 124814. doi:10.1016/j.biortech.2021.124814

Zhao, H., Fang, Y., Wang, X., Zhao, L., Wang, J., and Li, Y. (2018). Increasing L-threonine production in *Escherichia coli* by engineering the glyoxylate shunt and the L-threonine biosynthesis pathway. *Appl. Microbiol. Biotechnol.* 102 (13), 5505–5518. doi:10.1007/s00253-018-9024-3

Zhao, L., Lu, Y., Yang, J., Fang, Y., Zhu, L., Ding, Z., et al. (2020). Expression regulation of multiple key genes to improve L-threonine in *Escherichia coli*. *Microb. Cell Fact.* 19 (1), 46. doi:10.1186/s12934-020-01312-5

Zhu, L., Fang, Y., Ding, Z., Zhang, S., and Wang, X. (2019). Developing an l-threonine-producing strain from wild-type *Escherichia coli* by modifying the glucose uptake, glyoxylate shunt, and l-threonine biosynthetic pathway. *Biotechnol. Appl. Biochem.* 66 (6), 962–976. doi:10.1002/bab.1813

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 31 May 2023
DOI 10.3389/fbioe.2023.1188119

OPEN ACCESS

EDITED BY
Yun Chen,
Chalmers University of Technology, Sweden

REVIEWED BY
Jin Hou,
Shandong University, China
Dongming Xie,
University of Massachusetts Lowell, United States

*CORRESPONDENCE
Wei Wen Su,
wsu@hawaii.edu

RECEIVED 16 March 2023
ACCEPTED 22 May 2023
PUBLISHED 31 May 2023

CITATION
Han Z, Maruwan J, Tang Y and Su WW (2023), Conditional protein degradation in *Yarrowia lipolytica* using the auxin-inducible degron.
Front. Bioeng. Biotechnol. 11:1188119.
doi: 10.3389/fbioe.2023.1188119

Conditional protein degradation in *Yarrowia lipolytica* using the auxin-inducible degron

Zhenlin Han[1], Jessica Maruwan[1], Yinjie Tang[2] and Wei Wen Su[1]*

[1]Department of Molecular Biosciences and Bioengineering, University of Hawai'i at Manoa, Honolulu, HI, United States, [2]Department of Energy, Environmental and Chemical Engineering, Washington University, Saint Louis, MO, United States

Conditional protein degradation is a powerful tool for controlled protein knockdown. The auxin-inducible degron (AID) technology uses a plant auxin to induce depletion of degron-tagged proteins, and it has been shown to be functional in several non-plant eukaryotes. In this study, we demonstrated AID-based protein knockdown in an industrially important oleaginous yeast *Yarrowia lipolytica*. Using the mini-IAA7 (mIAA7) degron derived from *Arabidopsis* IAA7, coupled with an *Oryza sativa* TIR1 (OsTIR1) plant auxin receptor F-box protein (expressed from the copper-inducible MT2 promoter), C-terminal degron-tagged superfolder GFP could be degraded in *Yarrowia lipolytica* upon addition of copper and the synthetic auxin 1-Naphthaleneacetic acid (NAA). However, leaky degradation of the degron-tagged GFP in the absence of NAA was also noted. This NAA-independent degradation was largely eliminated by replacing the wild-type OsTIR1 and NAA with the OsTIR1^{F74A} variant and the auxin derivative 5-Ad-IAA, respectively. Degradation of the degron-tagged GFP was rapid and efficient. However, Western blot analysis revealed cellular proteolytic cleavage within the mIAA7 degron sequence, leading to the production of a GFP sub-population lacking an intact degron. The utility of the mIAA7/OsTIR1^{F74A} system was further explored in controlled degradation of a metabolic enzyme, β-carotene ketolase, which converts β-carotene to canthaxanthin via echinenone. This enzyme was tagged with the mIAA7 degron and expressed in a β-carotene producing *Y. lipolytica* strain that also expressed OsTIR1^{F74A} controlled by the MT2 promoter. By adding copper and 5-Ad-IAA at the time of culture inoculation, canthaxanthin production was found to be reduced by about 50% on day five compared to the control culture without adding 5-Ad-IAA. This is the first report that demonstrates the efficacy of the AID system in *Y. lipolytica*. Further improvement of AID-based protein knockdown in *Y. lipolytica* may be achieved by preventing proteolytic removal of the mIAA7 degron tag.

KEYWORDS

conditional protein degradation, degron, metabolic engineering, synthetic biology, *Yarrowia lipolytica*

1 Introduction

Degron-mediated protein degradation is an important tool for rewiring metabolic pathways, studying protein functions, and creating novel synthetic-biology systems (Natsume and Kanemaki, 2017). Auxin-inducible degron (AID) is a degron system originating from plants, but it has been successfully applied to various non-plant

eukaryotic organisms including baker's yeast *Saccharomyces cerevisiae,* mammalian cells, and transgenic mice (Nishimura et al., 2009; Yesbolatova et al., 2020). It has gained considerable interests for biotechnological applications owing to its specificity, degradation speed, and inducibility.

In plants, the Transport Inhibitor Response 1 (TIR1) auxin-receptor F-box protein, which is a component of the SCF (Skp1, Cullins, F-box proteins) multi-subunit E3 ubiquitin ligase complex, recruits the auxin-responsive proteins in the presence of indole-3-acetic acid (IAA), for ubiquitination and degradation. It was discovered by Nishimura et al. (2009) that the auxin-dependent protein degradation pathway could be transferred to *S. cerevisiae* and mammalian cells, while the heterologous AID system consists of three main components: an AID degron (a motif derived from the auxin-responsive proteins, such as *Arabidopsis thaliana* IAA17) fused to a target protein to be degraded, the TIR1 protein, and the auxin inducer molecule. It was reasoned that the heterologously expressed TIR1 protein can bind the host Skp1 adaptor protein to form a functional SCF ubiquitin ligase complex in the heterologous host, since the Skp1 protein is conserved in many species including yeast and humans (Zhang et al., 1995). Therefore, the plant's F-box protein in theory should be able to interact with different species' Skp1 protein, which means that AID system should be applicable to other eukaryotic species. Unfortunately, this might not be true in all cases. For example, Kanke et al. (2011) reported inefficient binding of *A. thaliana* TIR1 (AtTIR1) to endogenous Skp1 in fission yeast (*Schizosaccharomyces pombe*). Fusing *S. pombe* Skp1 to AtTIR1 was shown to increase the AID degradation efficiency in the fission yeast. In this case, the low affinity between AtTIR1 and fission yeast Skp1 is no longer an issue. However, when expressed from a strong promoter, the AtTIR1-Skp1 fusion protein was found to be toxic to *S. pombe* (Kanke et al., 2011).

The original AID degron system (AID1) has several other drawbacks as well, such as auxin independent degradation and requirement of high doses of auxin (Yamaguchi et al., 2019). More recently, an improved AID2 system was developed, which involves mutating Phe74 to either Ala or Gly (i.e., F74A or F74G) in *Oryza sativa* TIR1 (OsTIR1), and uses 5-Adamantyl-IAA (5-Ad-IAA) or 5-phenyl-IAA (5-Ph-IAA) as the inducer (Yesbolatova et al., 2020). The F74 A/G mutation in OsTIR1 would enlarge its auxin-binding pocket, so that the mutant shows a high binding affinity towards 5-Ad-IAA and 5-Ph-IAA, and a low affinity towards IAA (Nishimura et al., 2020; Zhang et al., 2022). As a result, the AID2 system is reported to have higher degradation efficiency, much lower leaky degradation, and require much lower auxin doses, and it has been shown to work in *S. cerevisiae,* chicken DT40 cells, and various vertebrate cell lines (Nishimura et al., 2020; Nishimura and Fukagawa, 2021; Watson et al., 2021; Zhang et al., 2022). However, the AID systems have shown varying degrees of effectiveness in different host organisms, and it has not been validated in the industrially important oleaginous yeast, *Y.lipolytica*, which is the focus of this study.

Y. lipolytica is an important non-model yeast widely considered as a promising industrial chassis for valorizing renewable carbon feedstocks to a wide variety of high-value chemicals (Niehus et al., 2018). As an oleaginous yeast, it is very efficient in *de novo* lipid biosynthesis, and can accumulate a high level of lipids in the cells using simple sugars as the carbon source. Therefore, *Y. lipolytica* is widely used for converting renewable sugar feedstocks into lipids as biofuel. It is also capable of utilizing lipids as the sole carbon source, and hence a promising microbial platform for valorizing renewable waste lipid feedstock (Li et al., 2020). Significant efforts have been made to engineer *Y. lipolytica* as a chassis organism for producing useful compounds. While some genetic tools are available for engineering *Y. lipolytica* metabolism, addition of the degron technology to the toolbox will significantly expand the possibilities of improving existing traits or creating novel traits in this important organism.

The objective of this study was to evaluate the auxin-inducible degron system in *Y. lipolytica*. Superfolder GFP (Pedelacq et al., 2006) tagged with the mini-IAA7 degron tag (Li et al., 2019) was used as a model protein to characterize the auxin-dependent degradation by co-expressing OsTIR1 vs. OsTIR1^{F74A}, respectively. The degron system was further validated by studying auxin-dependent degradation of a metabolic enzyme, the β-carotene ketolase, and its effect on carotenoid production in an engineered *Y. lipolytica* strain. The present study is significant as it lays the foundation for applying the conditional-degron technology to create novel and improved *Y. lipolytica* microbial cell factories.

2 Materials and methods

2.1 Strains and chemicals

Escherichia coli strain DH5α was used for plasmid manipulation and propagation. *S. cerevisiae* strain EBY100 (Boder and Wittrup, 1997) was used for assembly of DNA fragments based on yeast recombination-based cloning. *Y. lipolytica* Po1g (Nicaud et al., 2002) and β-carotene producing *Y. lipolytica* strain ST6057 (Kildegaard et al., 2017) were used for protein expression and carotenoid production, respectively. Synthetic auxin derivatives 5-Ad-IAA and 5-Ph-IAA were ordered from TCI America (Portland, OR) and R&D Systems (Minneapolis, MN), respectively. All other chemicals are analytical grades from Sigma Aldrich (St. Louis, MO).

2.2 Plasmid construction

Plasmid pEHT-G, which targets the intE_4 locus for insertion of the GFP expression cassette and the hygromycin selection marker cassette, is constructed as follows. The TEF_{intron} promoter (Tai and Stephanopoulos, 2013) fragment (*PrTEFin*), which is used to drive the GFP expression, was amplified with the forward primer C1TIF and reverse primer TINR using the plasmid pCFB4666 (Addgene #106144) as template (Holkenbrink et al., 2018). The *GFP* fragment was amplified with the forward primer TIGF and reverse primer SGPR using the plasmid sfGFP-pBAD (Addgene #54519) as template. Then *PrTEFin* and *GFP* fragments were assembled into AsiSI-linearized pCfB5219 plasmid (Addgene #106135) using NEBuilder HiFi DNA assembly (New England Biolabs, Ipswich, MA).

To construct plasmid pEHT-GI for expressing the GFP-mIAA7 fusion protein, the *GFP* fragment was amplified with the forward primer C1TIF and reverse primer SFGR using the plasmid sfGFP-pBAD as template. The *mIAA7* fragment was obtained with the forward primer IA7F and reverse primer IA7LR using the *Yarrowia* codon-optimized *mIAA7* fragment (synthesized by Twist Bioscience, San Francisco, CA) as template. The two fragments were then assembled into AsiSI-linearized pCfB5219 plasmid using NEBuilder HiFi DNA assembly.

TABLE 1 Major *Y. lipolytica* strains used in this study.

Strain name	Gene cassettes expressed[a]	Host
GI	*PrTEFin-sfGFP-mIAA7-TPex20*	Po1g
GI/WT	*PrTEFin-sfGFP-mIAA7-TPex20/PrMT2-OsTIR1-TXpr2*	Po1g
GI/F74A	*PrTEFin-sfGFP-mIAA7-TPex20/PrMT2-OsTIR1^{F74A}-TXpr2*	Po1g
G	*PrTEFin-sfGFP-TPex20*	Po1g
MG	*PrMT2-sfGFP-TXpr2*	ST6057
GIWH	*PrTEFin-sfGFP-mIAA7-Wps-H6-TPex20*	ST6057
GIWH/F74A	*PrTEFin-sfGFP-mIAA7-Wps-H6-TPex20/PrMT2-OsTIR1^{F74A}-TXpr2*	ST6057
WH	*PrTEFin-Wps-H6-TPex20*	ST6057

[a]*PrTEFin*, TEF_{intron} promoter; *PrMT2*, MT2 promoter; *TPex20*, Pex20 terminator; *TXpr2*, Xpr2 terminator; *sfGFP*, *mIAA7*, *Wps*, and *H6*, genes encoding sfGFP, mIAA7, Wps, and hexa-His tag, respectively.

To construct pEHT-GIWH, the *Wps-H6* fragment encoding the β-carotene ketolase from *Paracoccus* sp. N81106 was obtained with the forward primer A7WF and reverse primer PWH6R using the genomic DNA of ST7403 (Kildegaard et al., 2017) as template to introduce a hexa-His (H6) tag at the C-terminus of Wps. Subsequently, the *Wps-H6* fragment was assembled into SalI-linearized pEHT-GI using NEBuilder HiFi DNA assembly. To construct pEHT-WH, the *Wps-H6* fragment was amplified from the ST7403 genomic DNA using the forward primer TWPF and reverse primer PWH6R. The resulting product *Wps-H6* was then joined with *PrTEFin* and assembled into AsiSI-linearized plasmid pCfB5219 using NEBuilder HiFi DNA assembly.

To construct plasmids for integrating the OsTIR1 or OsTIR1^{F74A} expression cassette at the int_F2 locus, plasmid pFLM-G was first constructed by double digestion of pCU-IntF2U-LoxP-Leu2-LoxP-hp4D-XPR2-IntF2D (Li et al., 2020) with SalI and BamHI to remove the hp4D prompter fragment and assemble with the *MT2 promoter* (Xiong and Chen, 2020) and *GFP* fragments using NEBuilder HiFi DNA assembly. The *MT2 promoter* fragment was amplified using the forward primer MT2F and reverse primer MT2R, from the Po1g genomic DNA. The *GFP* fragment was amplified with the forward primer T2GF and reverse primer SFGN using plasmid sfGFP-pBAD as the template. The *OsTIR1* fragment was amplified with the forward primer MOSF and reverse primer MOSR, using a synthetic codon-optimized *OsTIR1* sequence as template. To obtain the F74A mutant of *OsTIR1*, the forward primer F74AF and reverse primer F74AR were used. Finally, *OsTIR1* and *OsTIR1^{F74A}* were assembled into BamHI/NheI linearized pFLM-G with NEBuilder HiFi DNA assembly to produce pFLM-OsTIR1 and pFLM-F74A, respectively. All primers used in this study are listed in Supplementary Table S1. Detailed plasmid maps and linear DNA structures of major constructs used in this study are presented in Supplementary Figures S1, S2.

2.3 Recombinant *Yarrowia* strain development

Plasmids pEHT-G and pEHT-GI were linearized with NotI and transformed into Po1g, and pFLM-G, and pEHT-GIWH were linearized with NotI and transformed into ST6057 using the lithium acetate (LiAc) method (Marsafari and Xu, 2020), with hygromycin B selection (250 μg/mL) on YPD plates. Resulting colonies were screened based on GFP fluorescence and further confirmed by PCR. Colony GFP fluorescence was visualized using the Dark Reader blue transilluminator (Clare Chemical Research, Dolores, CO). Plasmids pFLM-OsTIR1 and pFLM-F74A were linearized with NotI and NruI and then transformed into Po1g-EHT-GI, respectively. pFLM-F74A was also transformed into ST6057-EHT-GIWH and transformants selected using leucine dropout media. The resulting strains were screened with PCR, using genomic DNA as template. Major *Y. lipolytica* strains used in this study are summarized in Table 1.

2.4 Culture conditions for evaluating AID efficacy in *Y. lipolytica*

Po1g-GI/WT and Po1g-GI/F74A were grown at 28°C in 20 mL of YPD medium in a 250-mL baffled flask for 24 h, respectively, then $CuSO_4$ was added to a final concentration as specified. Twelve hours later, 1-Naphthaleneacetic acid (NAA) or 5-Ad-IAA (or 5-Ph-IAA, as specified) was added into Po1g-GI/WT and Po1g-GI/F74A cultures to a final concentration of 0.5 mM and 1 μM, respectively. The GFP fluorescence intensity and optical density of the culture were measured every hour to monitor the degradation of the sfGFP-mIAA7 fusion protein based on the GFP fluorescence and Western blot analysis. To evaluate the AID efficacy in regulating carotenoid biosynthesis in *Y. lipolytica*, ST6057-GIWH/F74A was cultured in YPD with 0.2 mM $CuSO_4$ and 1 μM 5-Ad-IAA added at the time of inoculation. All culture experiments were conducted in shake flasks.

2.5 Western blot analysis and GFP culture fluorescence measurement

All culture samples were centrifuged and cell pellets rinsed three times with PBS buffer and stored at −80°C for subsequent Western blot and/or culture GFP fluorescence analysis. Each cell pellet sample was thawed, resuspended in PBS buffer, and diluted to OD_{600} = 0.3 to measure the culture GFP fluorescence with a Hitachi F-2500 fluorescence spectrophotometer. To prepare

protein extracts for Western blot analysis, cell pellets were resuspended in 10% TCA buffer, followed by homogenization using Mini-Beadbeater-16 (Biospec model 607, Bartlesville, OK) with zirconia/silica beads (0.5 mm) in 3 × 1 min bursts. The extracted protein pellet was resuspended in a resuspension buffer as described previously (Cox et al., 1997). The protein concentration in the extract was measured using the Nanodrop ND-1000 spectrophotometer. The protein extract is then subject to SDS-PAGE and Western blot analysis. SDS-PAGE was performed using the 12% polyacrylamide gel; about 5 μg of total soluble protein for each sample was mixed with 5× loading buffer and subjected to electrophoresis. Proteins separated in SDS-PAGE gel were electroblotted onto a polyvinylidene difluoride (PVDF) membrane, and probed with anti-GFP antibody or anti-His Tag antibody (Genscript, Piscataway NJ), as described previously (Zhang et al., 2017).

2.6 qRT-PCR

ST6057-GI/F74A and ST6057-MG, respectively, was grown at 28°C in the YPD medium for 24 h, and $CuSO_4$ was then added into the culture at a final concentration of 0.2 mM. Cell culture samples were taken every 30 min for 3 h, and on the 4th hour. During each sampling, the cells were rinsed with sterile water, and the cell pellets were stored at −80°C. After all samples were collected, RNA was extracted from each sample with Quick Fungal Bacterial MiniPrep kit (ZymoResearch). The quality and concentration of RNA were assessed with Nanodrop ND-1000. Then RT reaction was conducted with 600 ng of RNA for each sample, using the LunaScript® RT SuperMix Kit (NEB). The resulting cDNA product was diluted 25-fold and used for qPCR with SYBR green dye. Primer sets Q74F1/Q74R1 and QSGF/QSGR were used in qPCR to detect *OsTIR1*^{F74A} and *GFP* transcripts, respectively. Primer sets QACT1F/QACT1R and QTEF1F/QTEF1R were used to detect transcript levels of ACT1 and TEF1 reference genes, respectively.

2.7 MT2 promoter-GFP fluorescence time course upon copper induction

ST6057-MG was grown at 28°C in the YPD medium for 24 h, then $CuSO_4$ was added into the culture at a final concentration of 0.2 mM. Cell culture samples were taken every 30 min for 3 h, and on the 4th hour. During each sampling, the cells were rinsed and resuspended in 200 μL of sterile water. Samples were diluted 20-fold and loaded into a 96-well plate in triplicates. Cell OD_{600} and culture GFP fluorescence were then measured with the Tecan Infinite M Plex plate reader. GFP fluorescence was measured with the excitation wavelength of 470 nm, and emission wavelength of 511 nm.

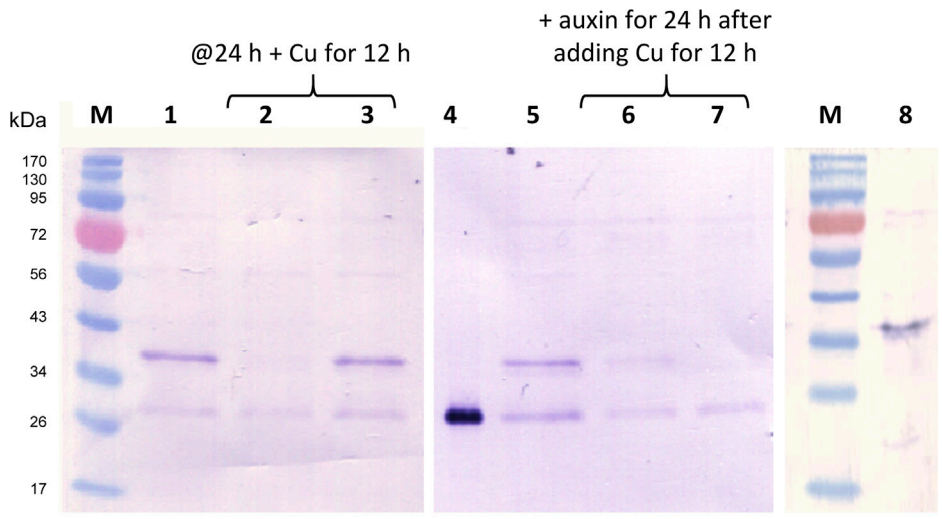
2.8 High-performance liquid chromatography (HPLC) analysis of carotenoids

ST6057-GIWH/F74A was grown at 28°C in the YPD medium. Three cultures were set up for comparison. At the time of inoculation, one culture was supplemented with $CuSO_4$ (0.2 mM), one with both $CuSO_4$ (0.2 mM) and 5-Ad-IAA (1 μM), and the third is a control, without supplementation of $CuSO_4$ or auxin. Samples were taken during the course of the culture to monitor the carotenoid production. The carotenoids in the cell samples were separated and quantified using HPLC as described previously (Li et al., 2020). The titer of the major carotenoid product, canthaxanthin, was estimated based on the calibration curve with a canthaxanthin standard (Sigma Aldrich 32993).

3 Results

3.1 Protein degradation using mIAA7-degron paired with OsTIR1/NAA vs. OsTIR1^{F74A}/5-Ad-IAA

To examine whether the AID system is functional in *Y. lipolytica*, a GFP reporter tagged with a C-terminal mIAA7 degron tag was expressed alone, or co-expressed with either the wild type OsTIR1 or its F74A variant (OsTIR1^{F74A}). We used superfolder GFP as the reporter instead of the enhanced GFP (EGFP) commonly used in *S. cerevisiae* because the latter was found in our prior studies to be inactive when expressed in *Y. lipolytica*. The mIAA7 degron is composed of amino acids 37–104 from the *A. thaliana* IAA7 protein (an auxin/IAA response transcription repressor). OsTIR1 is a plant auxin receptor F-box protein. Both OsTIR1 and OsTIR1^{F74A} genes were codon-optimized for *Y. lipolytica* expression. The expression of the GFP-mIAA7 fusion protein was driven by the TEF_{intron} promoter, whereas OsTIR1 (or OsTIR1^{F74A}) was expressed from the copper-inducible MT2 promoter (Xiong and Chen, 2020). NAA at 0.5 mM and 5-Ad-IAA at 1 μM were used with OsTIR1 and OsTIR1^{F74A}, respectively. Three *Y. lipolytica* strains GI, GI/WT, and GI/F74A were compared. The GI culture (which serves as a control) was inoculated into the YPD medium without adding additional chemicals. The GI/WT culture was inoculated into the YPD medium and cultured for 24 h, followed by supplementing with 0.2 mM of copper sulfate, and 12 h later 0.5 mM of NAA was added. The GI/F74A culture was treated similarly as the GI/WT culture, except that 0.5 mM NAA was replaced with 1 μM of 5-Ad-IAA. For all three cultures, samples were taken 36 h and 60 h post inoculation for Western blot analysis using an anti-GFP antibody. The result is presented in Figure 1. All samples in the Western blot were loaded based on the same total soluble protein concentration. For the GI control culture, besides the full-length fusion protein (with a molecular mass of 34.4 kDa), another immunoreactive band with a molecular mass similar to GFP was also detected on the Western blot. The proportion of this lower band grew overtime between 36 and 60 h post inoculation (cf. lanes 1 and 5 in Figure 1). This result suggests that the C-terminal mIAA7 tag may be susceptible to intracellular proteolytic cleavage in *Y. lipolytica*. For the GI/WT culture, the GFP-mIAA7 fusion protein was found to be degraded even with only Cu^{2+} (but without NAA) added (cf. lanes 2 and 8, Figure 1), and a very small amount of the fusion protein was detected 24 h after adding NAA (lane 6, Figure 1). This result indicated leaky degradation and inefficient NAA-specific protein degradation for the wildtype OsTIR1/NAA system. This finding was further



Lane	1	2	3	4	5	6	7	8
Strain	GI	GI/WT	GI/F74A	G	GI	GI/WT	GI/F74A	GI/WT
Time after inoculation (h)	36	36	36	-	60	60	60	36
Cu^{2+}	-	+	+	-	-	+	+	-
NAA or 5-Ad-IAA	-	-	-	-	-	+	+	-

FIGURE 1
GFP Western blot analysis of protein degradation using the mIAA7-degron paired with OsTIR1/NAA and $OsTIR1^{F74A}$/5-Ad-IAA, respectively. Refer to text for further experimental detail.

corroborated by comparing the GFP fluorescence of the GI/WT culture samples which shows a decrease in culture GFP fluorescence after adding copper but no further drop in GFP fluorescence was noted after adding NAA (data not shown). Unlike wildtype OsTIR1, the F74A variant of OsTIR1 provided highly efficient degradation of GFP-mIAA7 when induced using the synthetic IAA derivative 5-Ad-IAA, and showed little signs of leaky degradation. This can be seen clearly on Figure 1 by comparing lanes 3 and 7 that shows highly efficient 5-Ad-IAA-induced degradation of GFP-mIAA7, while little or no leaky degradation was detected in the absence of 5-Ad-IAA by comparing lanes 1 and 3.

3.2 Effect of copper concentration on the mIAA7/$OsTIR1^{F74A}$/5-Ad-IAA AID system

In designing the AID systems for this study, we controlled OsTIR1 and $OsTIR1^{F74A}$ expression using the Cu^{2+}-inducible MT2 promoter. By tuning the copper induction condition to control the level of OsTIR1 or $OsTIR1^{F74A}$, it may help to minimize leaky protein degradation in the absence of exogenous auxin addition. In the case of $OsTIR1^{F74A}$, we showed in Figure 1 that no leaky degradation was noted upon MT2 promoter induction using copper sulfate at 0.2 mM. However, the concentration of $OsTIR1^{F74A}$ relative to the concentration of the degron-tagged protein may affect the efficiency of the AID mediated protein degradation. We therefore investigated the effect of copper induction concentration on the extent of AID-mediated GFP-mIAA7 degradation. Three copper inducer concentrations were tested (0.05, 0.2, and 0.3 mM). The GI/F74A culture was grown in YPD for 24 h, then Cu^{2+} was added at 0.05, 0.2, and 0.3 mM, respectively. At 36 h post inoculation, 1 µM of 5-Ad-IAA was added and the cultures were monitored at 0, 3, 5, 7, and 24 h post 5-Ad-IAA induction. The samples were analyzed using GFP Western blot (Figure 2A), from which, protein degradation was found to occur very rapidly (cf. 0 and 3 h samples), yet similar levels of degradation were noted at all three Cu^{2+} concentration tested. Given that the culture was induced by Cu^{2+} for 12 h before adding 5-Ad-IAA, sufficient $OsTIR1^{F74A}$ might have already been accumulated even at the lowest Cu^{2+} induction concentration and thus was not limiting. After 5-Ad-IAA induction, nearly all full-length GFP-mIAA7 was depleted within 3 h regardless the Cu^{2+} concentration used to induce the MT2 promoter. Interestingly, the faint upper band detected in the samples at 3, 5, and 7 h post 5-Ad-IAA induction showed a slightly lower molecular mass than the full-length GFP-mIAA7 detected in the 0-h sample. This faint upper band was almost completely depleted at 24 h post IAA addition for all three Cu^{2+} concentrations tested. All samples contain a lower band (likely GFP) similar to that noted in Figure 1. The GI/F74A cultivation was repeated using 0.2 mM Cu^{2+} and 1 µM 5-Ad-IAA, and GFP culture fluorescence was monitored hourly for up to 7 h (Figure 2B). The culture fluorescence data (per culture OD) corroborate the Western blot result, and further inform the gradual degradation of GFP-mIAA7 during the first 3 h post 5-Ad-IAA induction, while the

10.3389/fbioe.2023.1188119

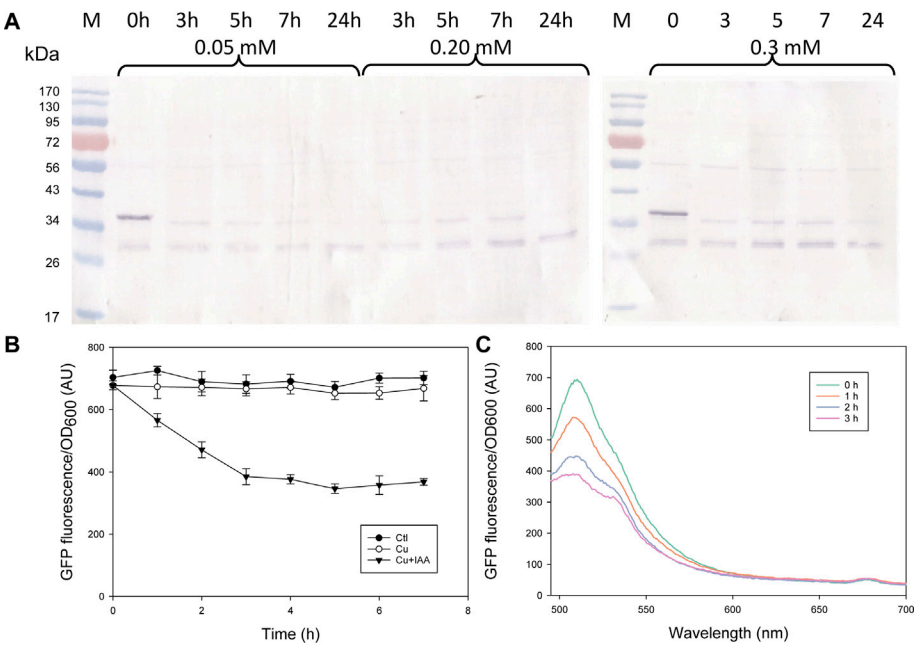


FIGURE 2
Characterization of the mIAA7/OsTIR1^{F74A}/5-Ad-IAA AID system: effect of Cu^{2+} induction concentrations (0.05, 0.2, vs. 0.3 mM) and degradation time course of GFP-mIAA7 in *Y. lipolytica*. **(A)** GFP Western blot analysis of samples taken at 0, 3, 5, 7, and 24 h upon addition of 1 μM of 5-Ad-IAA. GI/F74A culture was grown for 24 h, followed by Cu^{2+} (0.2 mM) induction (to express OsTIR1^{F74A}), and then 5-Ad-IAA was added 12 h later. **(B)** Time courses of GFP culture fluorescence upon adding Cu^{2+} (0.2 mM) alone or Cu^{2+} plus 5-Ad-IAA. **(C)** Changes in the GFP fluorescence spectra in response to addition of Cu^{2+} (0.2 mM) plus 5-Ad-IAA.

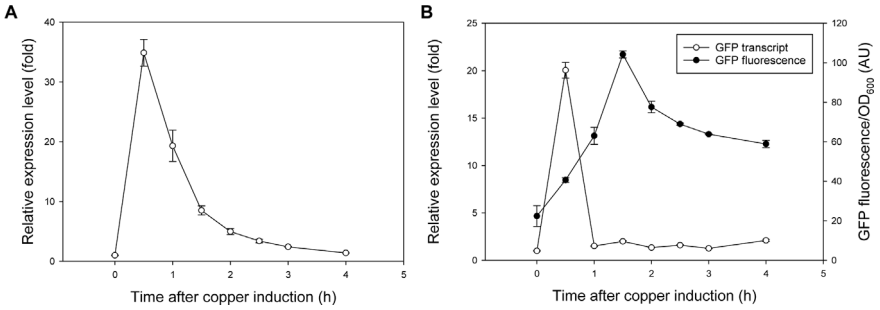


FIGURE 3
Induction kinetics of the MT2 promoter upon Cu^{2+} (0.2 mM) addition. **(A)** Transcript of OsTIR1^{F74A} expressed from the MT2 promoter (GI/F74A culture). **(B)** GFP expression driven by the MT2 promoter (MG culture), indicated by the GFP fluorescence and transcript time courses.

control and Cu^{2+}-only culture displayed essentially constant fluorescence throughout the duration of the test. The residual background fluorescence seen in Figure 2B resulted mainly from culture autofluorescence and to a less extent the GFP cleaved from GFP-mIAA7 that was not degraded by OsTIR1^{F74A}. As indicated in the fluorescence spectra of the culture samples (Figure 2C), the characteristic GFP emission peak occurring at around 510 nm subsided after 5-Ad-IAA addition, and essentially disappeared after 3 h.

Because GFP-mIAA7 was driven by the constitutive TEF_{intron} promoter, whereas OsTIR1^{F74A} was by the copper-inducible MT2 promoter, we tracked the *OsTIR1*^{F74A} transcript over a period of 4 h from the time of copper induction to inform the efficacy of the MT2 promoter (Figure 3A). The induction was found to be very rapid, and about 30~40-fold increase in the *OsTIR1*^{F74A} transcript was detected within 30 min after adding Cu^{2+} (0.2 mM), but the transcript level subsided quickly afterwards. We then examined another *Y. lipolytica* strain that expresses GFP driven by the MT2 promoter, by measuring time courses of GFP transcript and GFP fluorescence (indicating GFP protein concentrations) upon copper induction. As seen in Figure 3B, the *GFP* gene induction kinetics resembled that of the *OsTIR1*^{F74A} gene (also driven by the

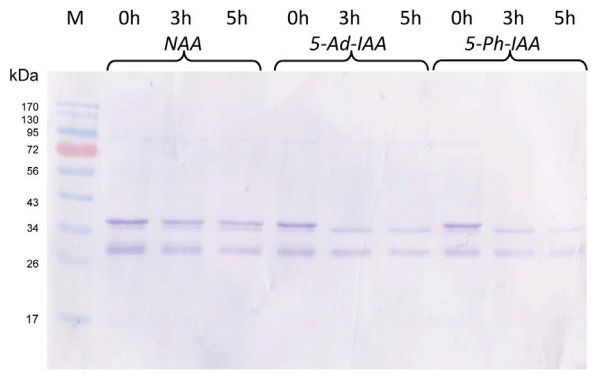


FIGURE 4
Characterization of the mIAA7/OsTIR1^{F74A} AID system: effect of different auxins. GFP Western blot analysis of samples taken at 0, 3, and 5 h upon addition of 1 μM of 5-Ad-IAA or 5-Ph-IAA, or 0.5 mM of NAA. The GI/F74A culture was grown for 24 h, followed by Cu^{2+} (0.2 mM) induction (to express OsTIR1^{F74A}), and then the auxin was added 12 h later.

MT2 promoter) in Figure 3A. Importantly, even though the *GFP* transcript quickly decreased to near the pre-induction level after it peaked, GFP protein accumulated above the pre-induction level. The OsTIR1^{F74A} protein level was not monitored, yet based on the data for GFP, and the fact that auxin-dependent GFP-mIAA7 degradation was found to be effective (Figures 1, 2), it is likely that the level of OsTIR1^{F74A} protein present in the *Y. lipolytica* cells was sufficient to enable efficient protein degradation.

3.3 Effect of auxin inducers on the mIAA7/OsTIR1^{F74A} AID system

Besides 5-Ad-IAA, use of 5-Ph-IAA with OsTIR1^{F74A} was also reported in several studies. To compare the effect of different auxin inducers on protein degradation, Po1g-GI/F74A was grown at 28°C in the YPD medium for 24 h, $CuSO_4$ added to a final concentration of 0.2 mM and further incubated for 12 more hours. NAA, 5-Ad-IAA, and 5-Ph-IAA were then supplemented at a final concentration of 0.5 mM, 1 μM, and 1 μM, respectively, and GFP-mIAA7 protein degradation was monitored using Western blot for up to 5 hours. As shown in Figure 4, 5-Ad-IAA and 5-Ph-IAA are equally effective in degrading the GFP-mIAA7. At 3 h post auxin induction, the full-length fusion protein was completely depleted, noting that the faint upper band seen under 3 and 5 h has a lower molecular mass than that of the full-length GFP-mIAA7 (as in Figure 2A). NAA on the other hand, despite at a much higher concentration, caused very modest protein degradation.

3.4 Conditional degradation of a biosynthetic enzyme using the mIAA7/OsTIR1^{F74A}/5-Ad-IAA AID system

The mIAA7/OsTIR1^{F74A} degron system was further validated by examining auxin-dependent degradation of β-carotene ketolase, and its effect on carotenoid production in an engineered *Y. lipolytica* strain. The host *Y. lipolytica* strain ST6057 was engineered to produce β-carotene (Kildegaard et al., 2017), and it was further engineered in this study to create ST6057-GIWH/F74A that overexpresses the β-carotene ketolase Wps to convert β-carotene to canthaxanthin. Three culture treatments were set up and compared. Cu^{2+} (0.2 mM) and 5-Ad-IAA (1 μM) or Cu^{2+} alone was introduced at inoculation, vs. a control culture without supplementing Cu^{2+} or 5-Ad-IAA. Each culture was allowed to grow for 5 days. On day 2 and day 5, cells were extracted for carotenoid analysis using HPLC, and the resulting chromatograms (based on the same amount of cell biomass extracted across all three culture treatments) are presented in Figure 5. In our previous study, we identified the carotenoids produced by the *Y. lipolytica* ST7403 strain (which was derived from ST6057 by overexpressing β-carotene ketolase and β-carotene hydroxylase) using triple quadrupole LC/MS and published mass to charge ratio (m/z) of known carotenoid species (Li et al., 2020). In Figure 5, major carotenoid species appeared in the HPLC chromatograms were identified as follows: 1) canthaxanthin (retention time: 11.5–12 min), 2) echinenone variants (retention time: 14–15.5 min), and 3) β-carotene (retention time: 24–24.5 min). The canthaxanthin titer in the culture received both Cu^{2+} and 5-Ad-IAA (5.32 ± 0.25 mg/L on day 2 and 6.39 ± 0.43 on day 5) was about half of that seen in the control (10.62 ± 0.39 on day 2, and 13.61 ± 0.87 mg/L on day 5) or the culture receiving Cu^{2+} alone (8.12 ± 0.26 mg/L on day 2 and 13.57 ± 0.77 mg/L on day 5). Before β-carotene was converted to canthaxanthin, it was first transformed into echinenone variants. On day 5, in the culture treated with Cu^{2+} and 5-Ad-IAA, the early echinenone variant (appeared between 15 and 15.5 min) accumulated to a level that was about twice as much as those seen in the other two cultures, whereas the later-stage echinenone variants (between 14 and 15 min) accumulated to a level that was about half of those seen in the other two cultures. Furthermore, β-carotene accumulation was detected only in the culture received Cu^{2+} and 5-Ad-IAA due to lower Wps activity that reduced further conversion of β-carotene. All of these data indicate that the mIAA7/OsTIR1^{F74A} degron system rendered a lower Wps activity by lowering its abundance in the cells via protein degradation. Residual ketolase activities however persisted, and led to the formation of canthaxanthin and echinenone variants. In a related test, Cu^{2+} and 5-Ad-IAA were added 1 day after cell inoculation, and a similar carotenoid product trend emerged, i.e., lower canthaxanthin/later-stage echinenone and higher β-carotene/early-stage echinenone (data not shown). Addition of Cu^{2+} alone had almost no effect on the carotenoid production compared to control, indicating no IAA-independent leaky protein degradation with the mIAA7/OsTIR1^{F74A} degron system.

To further investigate the system at the protein level, Western blot analysis was conducted (Figure 6). The theoretical molecular mass of GFP-mIAA7-Wps-H6 and Wps-H6 are 61.8 kDa and 27.8 kDa, respectively. From the anti-His-tag Western blot in Figure 6 (left panel), the control culture showed an upper band corresponding to the full-length fusion protein and a lower band with a size similar to that of Wps, but the culture with copper plus 5-Ad-IAA only showed the lower band, indicating essentially complete degradation of the degron-tagged full-length fusion protein in the cells. However, the cleaved Wps product in the cells might contain

10.3389/fbioe.2023.1188119

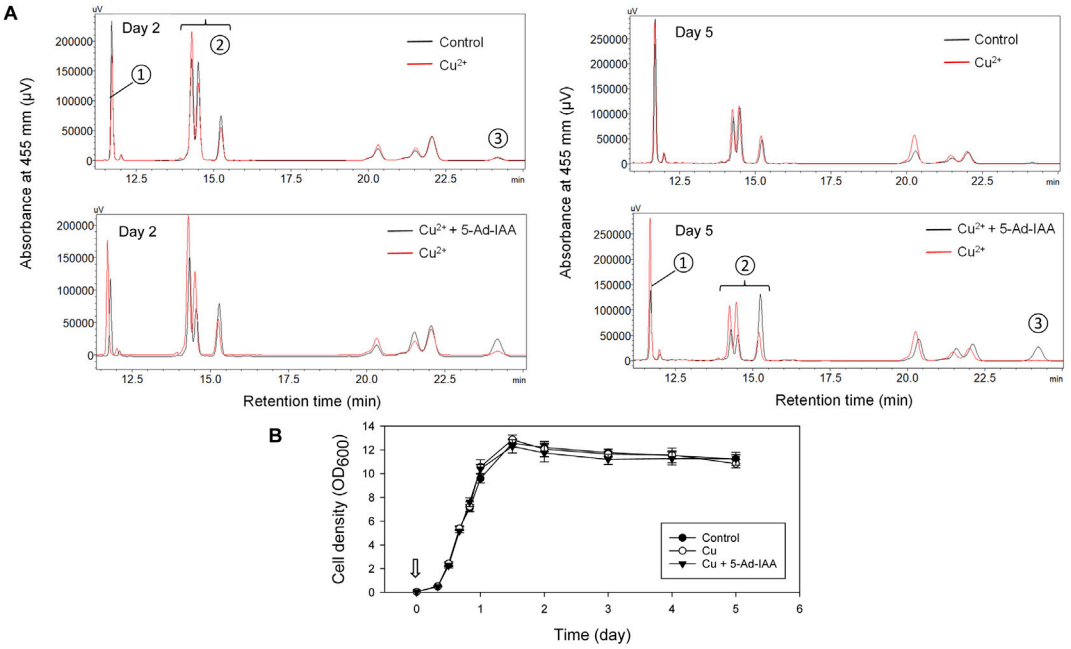


FIGURE 5
Controlling carotenoid biosynthesis in *Y. lipolytica* using the mIAA7/OsTIR1^{F74A} AID system as shown in HPLC analysis. Cu^{2+} (0.2 mM) and 5-Ad-IAA (1 μM) or Cu^{2+} alone was introduced at inoculation to the GIWH/F74A culture, and allowed to grow for 5 days. Control culture received no Cu^{2+} or 5-Ad-IAA. **(A)** HPLC analysis of samples taken on day 2 (left panel) and day 5 (right panel). Carotenoid species are labeled as follows: ① canthaxanthin, ② echinenone variants, and ③ β-carotene. **(B)** The corresponding growth curves (the arrow indicates the time Cu^{2+} and 5-Ad-IAA were added to the culture).

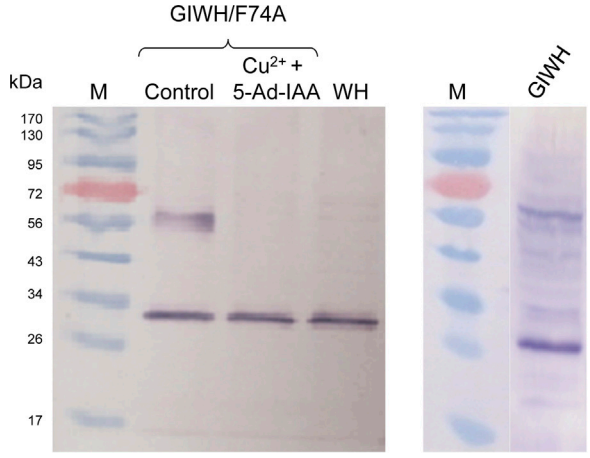


FIGURE 6
Western blot analysis of conditional degradation of β-carotene ketolase (Wps) using the mIAA7/OsTIR1^{F74A} AID system. Cu^{2+} (0.2 mM) and 5-Ad-IAA (1 μM) were introduced at inoculation to the GIWH/F74A culture and allowed to grow for 5 days before the cells were extracted for Western blot analysis using an anti-His-tag antibody (left panel). Control and WH cultures received no Cu^{2+} or 5-Ad-IAA. In a separate experiment, the GIWH culture (without co-expressing OsTIR1^{F74A}) was grown for 1 day, before the cells were extracted for Western blot analysis using an anti-GFP antibody (right panel).

no or only partial mIAA7 sequence which was insufficient for binding with OsTIR1^{F74A} and hence could not be degraded. The residual Wps was thus still able to convert β-carotene into downstream carotenoid products, and this may account for the canthaxanthin and echinenone variants seen in the HPLC chromatograms. In a separate experiment, *Y. lipolytica* strain ST6057-GIWH (without co-expressing OsTIR1^{F74A}) was grown for 1 day, before the cells were extracted for Western blot analysis using an anti-GFP antibody (Figure 6, right panel). A major cleaved product with size similar to GFP is clearly visible on the blot along with the full-length fusion protein. Multiple faint immunoreactive bands are also visible with sizes between those of the full-length protein and the cleaved GFP protein. This result indicates that proteolytic cleavage within the mIAA7 sequence noted above with the GFP-mIAA7 protein also occurred in the GFP-mIAA7-Wps-H6 protein, despite that the mIAA7 sequence is not present at the protein termini. The mIAA7 peptide consists of the conserved F-box protein binding motif flanked by disordered sequences. It is known that long and unstructured peptide linkers may be prone to proteolytic digestion (Chen et al., 2013). Therefore, following initial intracellular proteolytic cleavage within the mIAA7 sequence, the remnant mIAA7 peptides may be further digested by carboxyl and/or amino-peptidases, leading to the formation of the GFP and Wps-H6 products seen on the western blots (Figure 6).

4 Discussion

Since Nishimura et al. (2009) reported that the plant auxin-dependent protein degradation pathway could be transferred to non-plant eukaryotic cells, the AID system has been applied to several different hosts, yet its utility in *Y. lipolytica* had not been demonstrated prior to the present study. In this study, we choose to focus on the mIAA7 degron, instead of the more commonly used mAID or AID* (a shorter version of mAID) degrons. In *Caenorhabditis elegans*, higher protein degradation efficiency was seen with mIAA7 than with AID* (Sepers et al., 2022). Though derived from different *A. thaliana* IAA response transcription repressor proteins (IAA7 vs. IAA17), and differ in primary sequences, both mIAA7 (IAA7 37-104) and AID* (IAA17 71-114) degrons contain the conserved domain II F-box protein binding motif, yet the former has a longer amino-terminal extension (which was reported to be important to TIR1-mediated protein degradation in plants) and contains no putative ubiquitination sites (and hence the degron tag itself is not ubiquitinated) (Sepers et al., 2022). The mIAA7 degron was first reported by Li et al. (2019) and was shown to be an optimal degron to pair with the *A. thaliana* AFB2 F-box protein to enable protein degradation. Sepers et al. (2022) used mIAA7 in combination with AtOsTIR1^{F79G} in their AID system. In the present study, we showed that mIAA7 worked well in combination with OsTIR1^{F74A} and either 5-Ad-IAA or 5-Ph-IAA in inducing rapid auxin-dependent protein degradation in *Y. lipolytica*. Importantly, the mIAA7/OsTIR1^{F74A}/5-Ad-IAA degron system was shown in this study to be able to degrade both cytosolic (GFP) and integral membrane (Wps) proteins.

Whether it was fused internally or to the carboxyl terminus of the target protein, the mIAA7 degron was shown in this study to direct protein degradation in *Y. lipolytica*. However, Western blot analysis revealed cellular proteolytic cleavage within the mIAA7 degron sequence, leading to the production of target-protein subpopulations lacking an intact degron, which prevented complete degradation of the target protein upon auxin addition. By resolving the *A. thaliana* TIR1-auxin-IAA7 complex topology, it was shown that regions in the vicinity of the mostly conserved degron (VGWPP-[VI]-[RG]-x (2)-R) motif of IAA proteins are intrinsically disordered and they cooperatively position IAA protein on TIR1 (Niemeyer et al., 2020). The mIAA7 degron tag is only a portion of the IAA7 protein, and it consists of the conserved degron motif flanked by disordered sequences. The mIAA7 tag when fused between two other protein moieties or to the protein termini would likely be present as a highly disordered linker especially in the absence of the auxin inducer. Such a long (68 residues, 7.53 kDa) and disordered linker may be very prone to proteolytic digestion (Chen et al., 2013), which may explain why a portion of the degron-tagged proteins appeared to lose their degron tag, as shown in the Western blot results.

To reduce leaky degradation, in this study, the TIR1 expression was under the regulation of the copper-inducible MT2 promoter. As shown in Figure 3, MT2 promoter induction was very rapid, yet the transcript level subsided quickly after peaking at about 30 min after induction. The sharp decline in transcript level could be due to the copper (II) detoxification mechanism (Peng et al., 2015). Copper is an essential trace element yet it becomes toxic if not properly regulated. In the yeast *S. cerevisiae*, this mechanism entails reduction of copper (II) to copper (I), which is then bound to metallothionein or converted to copper metal, and may result in depletion of the cellular copper (II) inducer pool (Hassett and Kosman, 1995). A similar copper detoxification mechanism also exists in *Y. lipolytica* (Ran et al., 2023). Besides the copper-inducible promoter systems, alternative promoters (Sun et al., 2022) inducible by erythritol (Trassaert et al., 2017), xylose (Wei et al., 2020), and oleic acid (Sassi et al., 2016), respectively, may be considered for driving the TIR1 expression.

Compared with its F74A variant, the wild-type OsTIR1 performed poorly in *Y. lipolytica*, showing serious leaky degradation of mIAA7-tagged GFP in the absence of exogenous NAA addition. Auxin-independent leaky protein degradation with the wild-type OsTIR1 has been widely reported in several host systems (Yesbolatova et al., 2020). This phenomenon is generally believed (Mendoza-Ochoa et al., 2019) to result from 1) an intrinsic low affinity between the wild-type OsTIR1 and its substrate (i.e., the degron-tagged protein) even in the absence of auxin (Tan et al., 2007), 2) small amounts of auxin in the culture media, or 3) low levels of endogenous auxin in plants and yeast species such as *S. cerevisiae* (Rao et al., 2010) and *Y. lipolytica* (Gul Jan et al., 2019). The F74A mutation in OSTIR1 would enlarge its auxin-binding pocket to enable binding of 5-Ad-IAA or 5-Ph-IAA with a very high affinity, while its affinity for IAA is much lower (Nishimura et al., 2020; Zhang et al., 2022). As shown in Figure 4, when compared with 5-Ad-IAA, NAA induced much less protein degradation in combination with OsTIR1^{F74A} in *Y. lipolytica*.

In this study, β-carotene ketolase (Wps, encoded by *crtW*) was chosen as a target enzyme to illustrate the degron application in regulating metabolic pathways. This enzyme catalyzes the conversion of β-carotene to canthaxanthin (Supplementary Figure S3). By tagging the enzyme with the mIAA7 degron, we demonstrated that this enzyme could be degraded upon addition of Cu^{2+} and 5-Ad-IAA, and resulted in lower canthaxanthin production (Figure 5). The AID system enables post-translational regulation of proteins/enzymes via conditional degradation, which is complementary to transcriptional regulations such as use of inducible/repressible promoters and CRISPR activation (CRISPRa) or interference (CRISPRi), and post-transcriptional regulations such as RNA interference (RNAi). The AID approach is especially useful when the targeted genes are essential (and hence cannot be knocked out) and/or cannot be regulated transcriptionally. The inducible nature of the AID system and its rapid induction kinetics make it a powerful molecular tool for exerting tight temporal regulation of metabolic networks. The AID approach does have one drawback which is the need to tag the target endogenous proteins with the degron sequence for degradation, which requires modifying the host genome. However, with the advances in CRISPR genome editing, this barrier can be readily overcome. To exemplify its applications in metabolic engineering and synthetic biology, one may apply the AID in metabolic perturbation to elucidate metabolic network regulations, to redirect metabolic fluxes by cutting off byproduct synthesis, to implement temporal control of metabolic reaction networks, or to alter growth patterns (e.g., decoupling growth from product formation) by triggering depletion of protein targets essential for cell proliferation at a desired time point during the culture cycle. To this end, utility of AID1 (wt OsTIR1 coupled with the AID* degron tag, and NAA as the

auxin inducer) in metabolic engineering of *S. cerevisiae* was demonstrated by the Vickers' group who showed that AID1-mediated degradation of farnesyl pyrophosphate synthase increased the geranyl pyrophosphate pool which was redirected towards monoterpene production, whereas depleting acetyl-CoA carboxylase enabled decoupling of growth and production (Lu et al., 2021). As an extension of the present study to improve carotenoid production in *Y. lipolytica*, one may consider attenuating ergosterol synthesis via auxin-inducible degradation of squalene synthase as the cells enter stationary phase, to direct more farnesyl pyrophosphate towards geranylgeranyl pyrophosphate and downstream carotenoid biosynthesis.

5 Conclusion

In this study we demonstrated that the mIAA7/OsTIR1^{F74A}/5-Ad-IAA AID system is functional in the industrially important oleaginous yeast *Y. lipolytica*. Conversely, the mIAA7/OsTIR1/NAA system works poorly with considerable leaky auxin-independent protein degradation. Meanwhile, our Western blot analyses revealed some degrees of proteolytic cleavage within the mIAA7 degron sequence whether it was fused internally or to the carboxyl terminus of the target protein. Work is currently underway to resolve this issue by creating alternative protein scaffolds to stabilize the degron structure. Having established an effective conditional-degron system for *Y. lipolytica* will greatly expand the synthetic-biology toolbox for this important organism to develop novel and more advanced traits.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

ZH and JM conducted experiments, analysed data, and contributed to writing of the manuscript. YT reviewed/edited the manuscript and secured the funding. WS conceived and designed research, analysed data, wrote the manuscript, and secured the funding. All authors contributed to the article and approved the submitted version.

Funding

This work was supported in part by AFRI (award number 2020-67022-31146) from the USDA NIFA, and by the NIFA hatch project HAW05040-H and multistate project HAW05041-R.

Acknowledgments

The authors are grateful to Dr. Irina Borodina of the Technical University of Denmark for providing the *Y. lipolytica* ST6057 strain.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1188119/full#supplementary-material

References

Boder, E. T., and Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* 15, 553–557. doi:10.1038/nbt0697-553

Chen, X., Zaro, J. L., and Shen, W. C. (2013). Fusion protein linkers: Property, design and functionality. *Adv. Drug Deliv. Rev.* 65, 1357–1369. doi:10.1016/j.addr.2012.09.039

Cox, J. S., Chapman, R. E., and Walter, P. (1997). The unfolded protein response coordinates the production of endoplasmic reticulum protein and endoplasmic reticulum membrane. *Mol. Biol. Cell* 8, 1805–1814. doi:10.1091/mbc.8.9.1805

Gul Jan, F., Hamayun, M., Hussain, A., Jan, G., Iqbal, A., Khan, A., et al. (2019). An endophytic isolate of the fungus *Yarrowia lipolytica* produces metabolites that ameliorate the negative impact of salt stress on the physiology of maize. *BMC Microbiol.* 19, 3. doi:10.1186/s12866-018-1374-6

Hassett, R., and Kosman, D. J. (1995). Evidence for Cu(II) reduction as a component of copper uptake by *Saccharomyces cerevisiae*. *J. Biol. Chem.* 270, 128–134. doi:10.1074/jbc.270.1.128

Holkenbrink, C., Dam, M. I., Kildegaard, K. R., Beder, J., Dahlin, J., Domenech Belda, D., et al. (2018). EasyCloneYALI: CRISPR/Cas9-Based synthetic toolbox for engineering of the yeast Yarrowia lipolytica. *Biotechnol. J.* 13, e1700543. doi:10.1002/biot.201700543

Kanke, M., Nishimura, K., Kanemaki, M., Kakimoto, T., Takahashi, T. S., Nakagawa, T., et al. (2011). Auxin-inducible protein depletion system in fission yeast. *BMC Cell Biol.* 12, 8. doi:10.1186/1471-2121-12-8

Kildegaard, K. R., Adiego-Perez, B., Domenech Belda, D., Khangura, J. K., Holkenbrink, C., and Borodina, I. (2017). Engineering of Yarrowia lipolytica for production of astaxanthin. *Synth. Syst. Biotechnol.* 2, 287–294. doi:10.1016/j.synbio.2017.10.002

Li, N., Han, Z., O'donnell, T. J., Kurasaki, R., Kajihara, L., Williams, P. G., et al. (2020). Production and excretion of astaxanthin by engineered *Yarrowia lipolytica* using plant oil as both the carbon source and the biocompatible extractant. *Appl. Microbiol. Biotechnol.* 104, 6977–6989. doi:10.1007/s00253-020-10753-2

Li, S., Prasanna, X., Salo, V. T., Vattulainen, I., and Ikonen, E. (2019). An efficient auxin-inducible degron system with low basal degradation in human cells. *Nat. Methods* 16, 866–869. doi:10.1038/s41592-019-0512-x

Lu, Z., Peng, B., Ebert, B. E., Dumsday, G., and Vickers, C. E. (2021). Auxin-mediated protein depletion for metabolic engineering in terpene-producing yeast. *Nat. Commun.* 12, 1051. doi:10.1038/s41467-021-21313-1

Marsafari, M., and Xu, P. (2020). Debottlenecking mevalonate pathway for antimalarial drug precursor amorphadiene biosynthesis in *Yarrowia lipolytica*. *Metab. Eng. Commun.* 10, e00121. doi:10.1016/j.mec.2019.e00121

Mendoza-Ochoa, G. I., Barrass, J. D., Terlouw, B. R., Maudlin, I. E., De Lucas, S., Sani, E., et al. (2019). A fast and tuneable auxin-inducible degron for depletion of target proteins in budding yeast. *Yeast* 36, 75–81. doi:10.1002/yea.3362

Natsume, T., and Kanemaki, M. T. (2017). Conditional degrons for controlling protein expression at the protein level. *Annu. Rev. Genet.* 51, 83–102. doi:10.1146/annurev-genet-120116-024656

Nicaud, J. M., Madzak, C., Van Den Broek, P., Gysler, C., Duboc, P., Niederberger, P., et al. (2002). Protein expression and secretion in the yeast Yarrowia lipolytica. *FEMS Yeast Res.* 2, 371–379. doi:10.1111/j.1567-1364.2002.tb00106.x

Niehus, X., Crutz-Le Coq, A. M., Sandoval, G., Nicaud, J. M., and Ledesma-Amaro, R. (2018). Engineering *Yarrowia lipolytica* to enhance lipid production from lignocellulosic materials. *Biotechnol. Biofuels* 11, 11. doi:10.1186/s13068-018-1010-6

Niemeyer, M., Moreno Castillo, E., Ihling, C. H., Iacobucci, C., Wilde, V., Hellmuth, A., et al. (2020). Flexibility of intrinsically disordered degrons in AUX/IAA proteins reinforces auxin co-receptor assemblies. *Nat. Commun.* 11, 2277. doi:10.1038/s41467-020-16147-2

Nishimura, K., and Fukagawa, T. (2021). A simple method to generate super-sensitive AID (ssAID)-based conditional knockouts using CRISPR-based gene knockout in various vertebrate cell lines. *Bio Protoc.* 11, e4092. doi:10.21769/bioprotoc.4092

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* 6, 917–922. doi:10.1038/nmeth.1401

Nishimura, K., Yamada, R., Hagihara, S., Iwasaki, R., Uchida, N., Kamura, T., et al. (2020). A super-sensitive auxin-inducible degron system with an engineered auxin-TIR1 pair. *Nucleic Acids Res.* 48, e108. doi:10.1093/nar/gkaa748

Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* 24, 79–88. doi:10.1038/nbt1172

Peng, B., Williams, T. C., Henry, M., Nielsen, L. K., and Vickers, C. E. (2015). Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: A comparison of yeast promoter activities. *Microb. Cell Fact.* 14, 91. doi:10.1186/s12934-015-0278-5

Ran, M., Zhao, G., Jiao, L., Gu, Z., Yang, K., Wang, L., et al. (2023). Copper ion mediates yeast-to-hypha transition in *Yarrowia lipolytica*. *J. Fungi (Basel)* 9, 249. doi:10.3390/jof9020249

Rao, R. P., Hunter, A., Kashpur, O., and Normanly, J. (2010). Aberrant synthesis of indole-3-acetic acid in *Saccharomyces cerevisiae* triggers morphogenic transition, a virulence trait of pathogenic fungi. *Genetics* 185, 211–220. doi:10.1534/genetics.109.112854

Sassi, H., Delvigne, F., Kar, T., Nicaud, J.-M., Coq, A.-M. C.-L., Steels, S., et al. (2016). Deciphering how LIP2 and POX2 promoters can optimally regulate recombinant protein production in the yeast Yarrowia lipolytica. *Microb. Cell Factories* 15, 159. doi:10.1186/s12934-016-0558-8

Sepers, J. J., Verstappen, N. H. M., Vo, A. A., Ragle, J. M., Ruijtenberg, S., Ward, J. D., et al. (2022). The mIAA7 degron improves auxin-mediated degradation in *Caenorhabditis elegans*. *G3 Genes|Genomes|Genetics* 12, jkac222. doi:10.1093/g3journal/jkac222

Sun, M.-L., Shi, T.-Q., Lin, L., Ledesma-Amaro, R., and Ji, X.-J. (2022). Advancing Yarrowia lipolytica as a superior biomanufacturing platform by tuning gene expression using promoter engineering. *Bioresour. Technol.* 347, 126717. doi:10.1016/j.biortech.2022.126717

Tai, M., and Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metab. Eng.* 15, 1–9. doi:10.1016/j.ymben.2012.08.007

Tan, X., Calderon-Villalobos, L. I. A., Sharon, M., Zheng, C., Robinson, C. V., Estelle, M., et al. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* 446, 640–645. doi:10.1038/nature05731

Trassaert, M., Vandermies, M., Carly, F., Denies, O., Thomas, S., Fickers, P., et al. (2017). New inducible promoter for gene expression and synthetic biology in Yarrowia lipolytica. *Microb. Cell Fact.* 16, 141. doi:10.1186/s12934-017-0755-0

Watson, A. T., Hassell-Hart, S., Spencer, J., and Carr, A. M. (2021). Rice (Oryza sativa) TIR1 and 5'adamantyl-IAA significantly improve the auxin-inducible degron system in *Schizosaccharomyces pombe*. *GenesGenes (Basel)* 12, 882. doi:10.3390/genes12060882

Wei, W., Shang, Y., Zhang, P., Liu, Y., You, D., Yin, B., et al. (2020). Engineering prokaryotic transcriptional activator XylR as a xylose-inducible Biosensor for transcription activation in yeast. *ACS Synth. Biol.* 9, 1022–1029. doi:10.1021/acssynbio.0c00122

Xiong, X., and Chen, S. (2020). Expanding toolbox for genes expression of *Yarrowia lipolytica* to include novel inducible, repressible, and hybrid promoters. *ACS Synth. Biol.* 9, 2208–2213. doi:10.1021/acssynbio.0c00243

Yamaguchi, N., Colak-Champollion, T., and Knaut, H. (2019). zGrad is a nanobody-based degron system that inactivates proteins in zebrafish. *Elife* 8, e43125. doi:10.7554/elife.43125

Yesbolatova, A., Saito, Y., Kitamoto, N., Makino-Itou, H., Ajima, R., Nakano, R., et al. (2020). The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat. Commun.* 11, 5701. doi:10.1038/s41467-020-19532-z

Zhang, B., Rapolu, M., Kumar, S., Gupta, M., Liang, Z., Han, Z., et al. (2017). Coordinated protein co-expression in plants by harnessing the synergy between an intein and a viral 2A peptide. *Plant Biotechnol. J.* 15, 718–728. doi:10.1111/pbi.12670

Zhang, H., Kobayashi, R., Galaktionov, K., and Beach, D. (1995). p19skp1 and p45skp2 are essential elements of the cyclin A-CDK2 S phase kinase. *Cell* 82, 915–925. doi:10.1016/0092-8674(95)90271-6

Zhang, X. R., Zhao, L., Suo, F., Gao, Y., Wu, Q., Qi, X., et al. (2022). An improved auxin-inducible degron system for fission yeast. *G3 (Bethesda)* 12, jkab393. doi:10.1093/g3journal/jkab393

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Methods
PUBLISHED 17 July 2023
DOI 10.3389/fbioe.2023.1202388

OPEN ACCESS

EDITED BY
Wei Luo,
Jiangnan University, China

REVIEWED BY
Quanfeng Liang,
Shandong University, China
Amir Pandi,
Université Paris Cité, France

*CORRESPONDENCE
Ramesh K. Jha,
rjha@lanl.gov

†These authors have contributed equally to this work

RECEIVED 08 April 2023
ACCEPTED 25 June 2023
PUBLISHED 17 July 2023

CITATION
Huttanus HM, Triola E-KH, Velasquez-Guzman JC, Shin S-M, Granja-Travez RS, Singh A, Dale T and Jha RK (2023), Targeted mutagenesis and high-throughput screening of diversified gene and promoter libraries for isolating gain-of-function mutations.
Front. Bioeng. Biotechnol. 11:1202388.
doi: 10.3389/fbioe.2023.1202388

Targeted mutagenesis and high-throughput screening of diversified gene and promoter libraries for isolating gain-of-function mutations

Herbert M. Huttanus[1,2†], Ellin-Kristina H. Triola[1,2†], Jeanette C. Velasquez-Guzman[1,2], Sang-Min Shin[1,3], Rommel S. Granja-Travez[1,3], Anmoldeep Singh[1], Taraka Dale[1,2,3] and Ramesh K. Jha[1,2,3]*

[1]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, United States, [2]Agile BioFoundry, Emeryville, CA, United States, [3]BOTTLE Consortium, Golden, CO, United States

Targeted mutagenesis of a promoter or gene is essential for attaining new functions in microbial and protein engineering efforts. In the burgeoning field of synthetic biology, heterologous genes are expressed in new host organisms. Similarly, natural or designed proteins are mutagenized at targeted positions and screened for gain-of-function mutations. Here, we describe methods to attain complete randomization or controlled mutations in promoters or genes. Combinatorial libraries of one hundred thousands to tens of millions of variants can be created using commercially synthesized oligonucleotides, simply by performing two rounds of polymerase chain reactions. With a suitably engineered reporter in a whole cell, these libraries can be screened rapidly by performing fluorescence-activated cell sorting (FACS). Within a few rounds of positive and negative sorting based on the response from the reporter, the library can rapidly converge to a few optimal or extremely rare variants with desired phenotypes. Library construction, transformation and sequence verification takes 6–9 days and requires only basic molecular biology lab experience. Screening the library by FACS takes 3–5 days and requires training for the specific cytometer used. Further steps after sorting, including colony picking, sequencing, verification, and characterization of individual clones may take longer, depending on number of clones and required experiments.

KEYWORDS

protein engineering, synthetic biology, promoter engineering, mutagenesis, polymerase chain reaction, overlap extension PCR, fluorescence-activated cell sorting, whole-cell biosensor

1 Introduction

In the field of synthetic biology, rational design of proteins and promoters has gained extensive interest, especially for metabolic engineering efforts (Blazeck and Alper, 2013; Xiong et al., 2021). For proteins, this is frequently achieved by site directed mutagenesis of specific codons in the genes (Alberghina and Lotti, 2005). For promoter engineering, the method is less streamlined, but commonly randomization of ribosomal binding sites

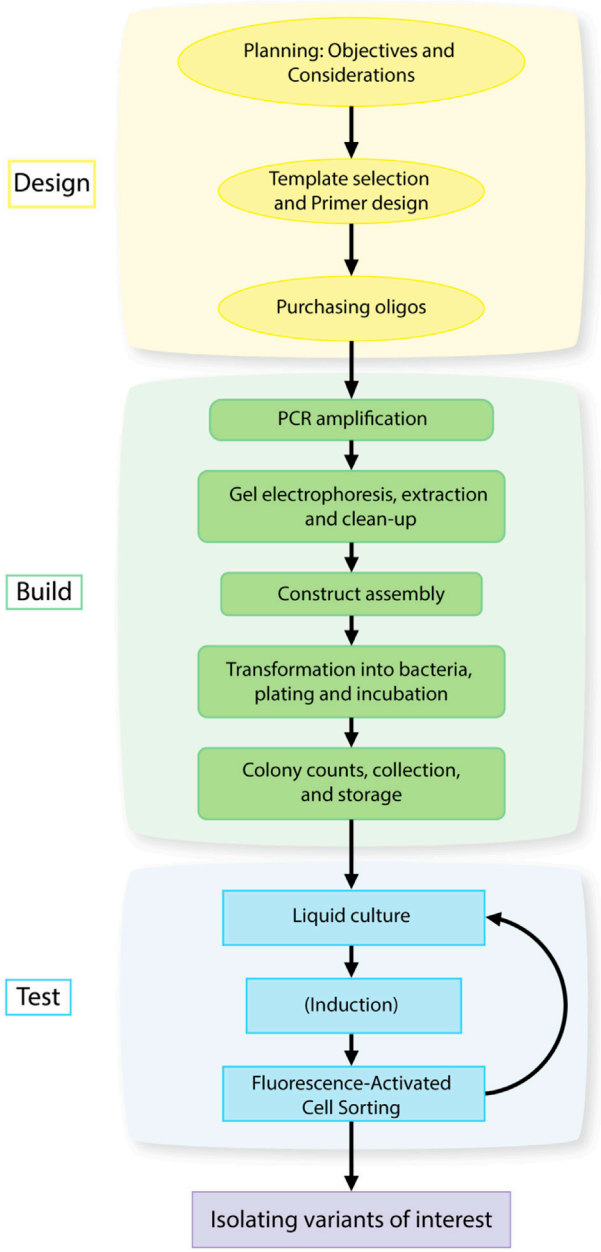


FIGURE 1
Design-Build-Test workflow for isolating gain-of-function mutations in a gene or promoter using library generation and high-throughput screening.

(RBS) is a preferred method for tuning the expression of genes (Salis et al., 2009; Zhang et al., 2015; Oesterle et al., 2017; Segall-Shapiro et al., 2018). Targeted mutagenesis is an essential method for achieving gain-of-function mutations in a gene or promoter. However, introducing single mutations one at a time and individually testing for changes in function is a tedious and time-intensive process, which at best only results in incremental changes to phenotype after each mutation. The protocol described herein eliminates the bottleneck of individually testing variants of genes and promoters (generated by site-directed mutagenesis) by assaying the combinatorial effect of several mutations at once, as part of a large multi-variant library.

The method described here uses a combination of overlap extension Polymerase Chain Reaction (PCR) (Ho et al., 1989; Horton et al., 1989) and saturation or partial saturation mutagenesis with degenerate primers (Kretz et al., 2004) to produce a library of gene and promoter variants that can then be screened for desired characteristics (outlined in Figure 1). Oligonucleotide overlap extension with degenerate codons is a simple yet powerful technique to introduce massive numbers of mutations, while using only a relatively simple two-step PCR. Economically, it takes advantage of the low cost for oligonucleotides (<70 bp), and the possibility of introducing degeneracy in a desired region. These oligos can then be used in a two-step PCR (fragment generation followed by assembly of fragments) to rapidly generate libraries with diversity on the order of 10^4–10^7 variants. The resulting libraries lend themselves to high-throughput screening via fluorescence activated cell sorting (FACS), when coupled to a relevant fluorescent reporter, allowing for the rapid identification and isolation of variants of interest, which can then be individually characterized.

For the successful application of this method, a variety of factors need to be taken into account when designing a library. The following sections outline important considerations when aiming to design, construct and screen a promoter or a protein library (Sections 1.1, 1.2 respectively).

1.1 Promoter and RBS libraries

Promoters may need to be modified for a variety of reasons including, tuning gene expression (Bakke et al., 2009; Boldrin et al., 2017), pathway optimization (Jin et al., 2019), designing synthetic circuits (Xie and Fussenegger, 2018), engineering biosensors (Pardo et al., 2020; Bentley et al., 2020) and providing new basic molecular biology tools for non-model organisms (Mordaka and Heap, 2018). While many of the mechanisms linking promoter sequence to transcription rates are known and several bioinformatic tools (Cassiano and Silva-Rocha, 2020; LaFleur et al., 2022) can help predict promoter strength from sequence alone, it is still necessary to test engineered promoters experimentally. The need for combinatorial building and testing of promoter regions and sequences in its proximity is commonly the method of choice. Thus, many attempts to modify promoter performance rely on semi-rational mutation libraries. In this context, semi-rational refers to the approach of targeting specific regions of the promoter, known to be involved with various mechanisms of transcription or translation, rather than just randomly mutating the entire promoter region.

For transcription, the regions approximately 35 and 10 bases upstream of the transcriptional initiation site (referred to as −35/−10 promoter sites) are particularly important for transcription initiation and therefore a single nucleotide mutation here can have dramatic effects (Einav and Phillips, 2019). Sometimes, if the goal is to create more subtle changes in translation rate or a series of promoters with steadily increasing or decreasing strength, it may be beneficial to mutate nearby areas, but leave the −35 and −10 regions unmutated (Mordaka and Heap, 2018). When importing a promoter

from one organism into another host, however, it becomes imperative to also change the −35/−10 sites. This strategy has contributed to achieving gain-of-function, allowing novel biosensors to be created in new host organisms such as *Pseudomonas putida* (Bentley et al., 2020), *Acinetobacter baylyi* ADP1 (Pardo et al., 2020) and *Corynebacterium glutamicum (Velasquez-Guzman and Huttanus et al, unpublished data).*

Regions around or within the −35/−10 sites can also contain operator regions for transcription factors. These operator regions are usually palindromic or pseudo-palindromic sequences to which the DNA binding domains of the transcription regulator bind. There are usually anywhere from one to three such sequences around the promoter region. Modification of the operator region can result in modulated binding affinity of the transcription regulator to the operator region and hence altered function. Randomization of only a few nucleotides in the operator region showed a wide range of repression levels of LacI that included increase in amplitude of response, very tight repression, or very weak repression resulting in constitutive activity of the promoter (Maity et al., 2012).

In addition to adjusting transcription rates via promoter and operator sites, further gene expression can be controlled at the translational level. While not technically part of the promoter, the ribosome binding site or RBS is typically located just between the promoter and the gene to be regulated, allowing for modifications to the RBS to be conveniently included in the promotor library. Mutation libraries of the RBS have been used to modulate and optimize translation rates in a variety of applications (Oesterle et al., 2017). Similarly, *cis*-acting elements on the mRNA can have a profound effect on translation rates (Gebauer et al., 2012; Rhodius et al., 2012) and have also been diversified to generate libraries with a wide range of expression levels of downstream genes (White, 2015; Pandey et al., 2022).

Through these and other methods, many sets of constitutive promoters covering a wide range of transcription rates have been developed. Yet there is still a need to engineer promoter sets for microbial hosts that include living therapeutics and other microbiota (Waller et al., 2017; Charbonneau et al., 2020; Dosoky et al., 2020), as well as non-model host strains for biomanufacturing purposes. A need for tuning constitutive promoters arises especially when the gene product shows instability in function or is toxic to the microbial host. In such cases, randomization of specific regions in the promoter such as −35/−10 sites, can tune down the constitutive promoter, resulting in stable expression of the downstream gene. The approach was successfully applied in tuning down the expression of *mucK* transporter gene for stable expression from a constitutive promoter in *Pseudomonas putida* (Shin et al., 2022).

Inducible promoters are an important component of the molecular biology toolkit (Chen et al., 2018), as they provide timely or dynamic regulation, and are often included in gene circuits (Mayo et al., 2006; Xie and Fussenegger, 2018) and metabolic pathway optimization (Jin et al., 2019). When engineering a synthetic inducible promoter, important aspects to consider include the inducibility and background activity of a promoter. Induction should be relatively straightforward for most cases, not requiring expensive chemicals, specific growth media or temperature shifts. In addition, an inducible promoter should show distinct activity when the inducer is present and low background activity when the inducer is absent. It is important to have tight regulation to avoid the basal expression levels interfering with the interpretation of the results (Hartman et al., 2011). Engineering these desirable qualities into a promoter, requires knowledge of the induction mechanism for targeted mutagenesis and, sometimes in addition, a semi-rational design of a library from which the desirable response can be isolated.

The mechanism of promoter regulation can vary and is determined by the type of transcription factor that interacts with those promoters. The most common ones are transcriptional repressors that bind to specific regions in the proximity of the promoter and block transcription (Figure 2A). Repression can be released by the repressor reacting to changes in the environment, such as binding to a specific ligand. Certain other promoters are regulated by transcriptional activators (Figure 2B) that recruit transcription machinery in response to a certain change in the environment, such as accumulation or depletion of a certain metabolite or change in temperature. Another very common kind of transcriptional regulator in bacteria works as a repressor as well as an activator. The LysR-Type Transcriptional Regulator (LTTR) (Maddocks and Oyston, 2008) forms homotetramers with two arms each bearing a pair of DNA binding domains (Figure 2C). Typically, one of those arms remains anchored to an operator site in proximity to the promoter, regardless of whether the protein is in the apo form or bound to a co-inducer. The other arm typically shifts its position from a second to a third operator site in response to conformational changes brought about by activation, such as binding to a co-inducer. These conformational shifts can control gene expression by exposing or occluding the −35/−10 regions of the promoter, by altering DNA bending, or by direct interaction with the RNA polymerase complex. Mutations to specific regions on the operators that are differentially bound in the active or repressed state can alter the dynamics of the conformational switch. This strategy has been used to develop a biosensor for *cis,cis*-muconic acid in *P. putida* using an LTTR CatM from *A. baylyi* ADP1 (Bentley et al., 2020).

In addition, a meaningful biosensor application relies heavily on the dynamic range or the maximal fold change in response over the basal levels. Even if a transcription regulator is found that is responsive to the molecule of interest, the dynamic range may need to be altered through further promoter engineering, typically using the approaches described above, as well as protein engineering of the transcriptional regulator itself (Jha et al., 2016; Jha et al., 2018; Bentley et al., 2020) (described below).

1.2 Protein libraries

Proteins may need to be modified for a range of reasons, such as altering ligand binding, DNA binding, stability, activity or protein-protein interactions. When targeting which sequences to randomize, one may take either a random, semi-rational or fully rational approach. This is frequently determined by the amount of information available for the protein of interest, and impacts the library size to pursue. Typically, saturation mutagenesis is only used to fully randomize a few positions in the protein sequence simultaneously, otherwise, the number of possible combinations of variants quickly becomes impractical to build individually in the laboratory or test in a given timeline, even with state-of-the-art technologies.

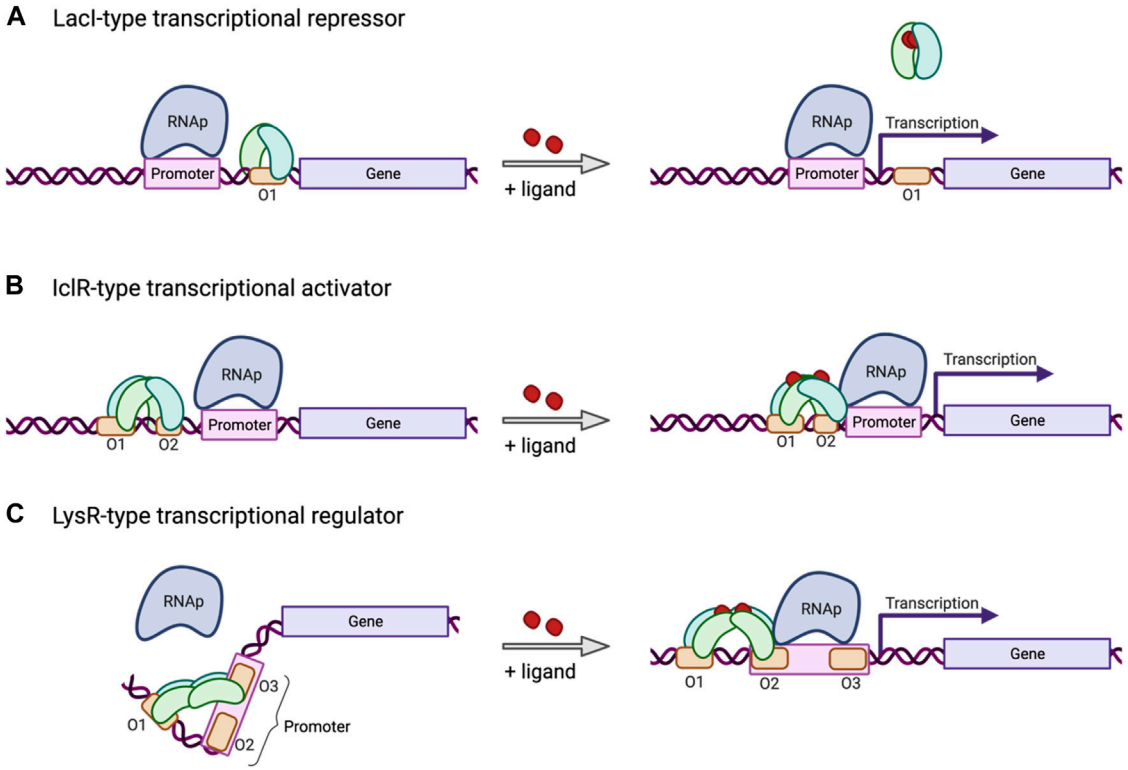


FIGURE 2
Mechanism of three common types of inducible transcriptional regulators. **(A)** The LacI-type transcriptional repressor binds an operator site located downstream from the promoter, or between the −10 and −35 sites, and blocks transcription, unless released by environmental factors, such as ligand binding. **(B)** The IclR-type transcriptional regulator induces transcription in response to environmental factors, such as ligand binding, inducing a conformational change in the tetramer bound to an operator site in proximity to the promoter. **(C)** The LysR-type transcriptional regulator (LTTR) binds operator sites 1 and 3 in its tetrameric apo form, bending the bound DNA and repressing transcription. In response to co-inducer binding, the LTTR tetramer will shift binding from operator site 3 to site 2, releasing the DNA bend, freeing up the promoter, and recruiting RNAp to allow transcription to be initiated. Created with BioRender.com.

Using rational targeted mutagenesis to alter the function, binding, or stability of a protein of interest is only a feasible approach if the structure and structure-function relationships of the protein are known. Ideally, the crystal structure of the protein is available to be pulled from a database such as PDB (Berman et al., 2000) or Uniprot (The UniProt Consortium Martin et al., 2021). Tertiary structures can also be predicted for proteins with a known sequence using computational approaches such as Rosetta (Rohl et al., 2004) or AlphaFold (Jumper et al., 2021). Further, the computational models facilitate docking of ligands in the putative binding sites, allowing for the visualization of residues to target for mutagenesis. *In silico* docking consisting of protein-protein or protein-ligand interactions (Lyskov and Gray, 2008; Lyskov et al., 2013) permits guided selection of amino acids involved in binding. Targeting only very relevant residues with computationally informed mutations, or conservative mutations such as neutral drift mutations (Lynch and Hill, 1986), can simplify library generation, reduce library size, and decrease noise and workload. *In silico* ligand docking in a comparative model to guide the design of a focused library was successfully used to engineer several biosensors for small molecules (Jha et al., 2015; Jha et al., 2016; Shin et al., 2022).

The ability to diversify multiple positions in a given sequence at once provides a significantly faster route to arriving at a combination of mutations providing appreciable alterations in phenotype. In addition, even imperfect predicted structures can be used to great effect by indicating promising residues to target for mutagenesis (Jha et al., 2016; Shin et al., 2022).

Once targeted positions in the genetic sequence have been identified, they are then mutagenized by PCR amplification and assembly using primers with incorporated degenerate codons, as shown in Figure 3. These degenerate codons have variation in the identity of the base at one or more positions such that the oligonucleotide pool contains unique sequences covering the range of mutations. Degenerate primers have long been used in the amplification and detection of panels of proteins, e.g., in virology (Li et al., 2012). When several degenerate primers targeting different regions of a gene are used for PCR fragments and assembly, the result is a combinatorial library consisting of all possible codon variations and combinations thereof as seen in Figure 4.

When designing the primers, it is helpful to refer to an amino acid substitution matrix, such as the Block Substitution Matrix 62 (BLOSUM62) table (Supplementary Table S1) (Henikoff and Henikoff, 1992). This 2-D matrix shows the log-odds score of finding two given amino acids in alignment, that is, comparing the occurrence of such an alignment to one that would be expected by random chance. A positive score indicates that this alignment is

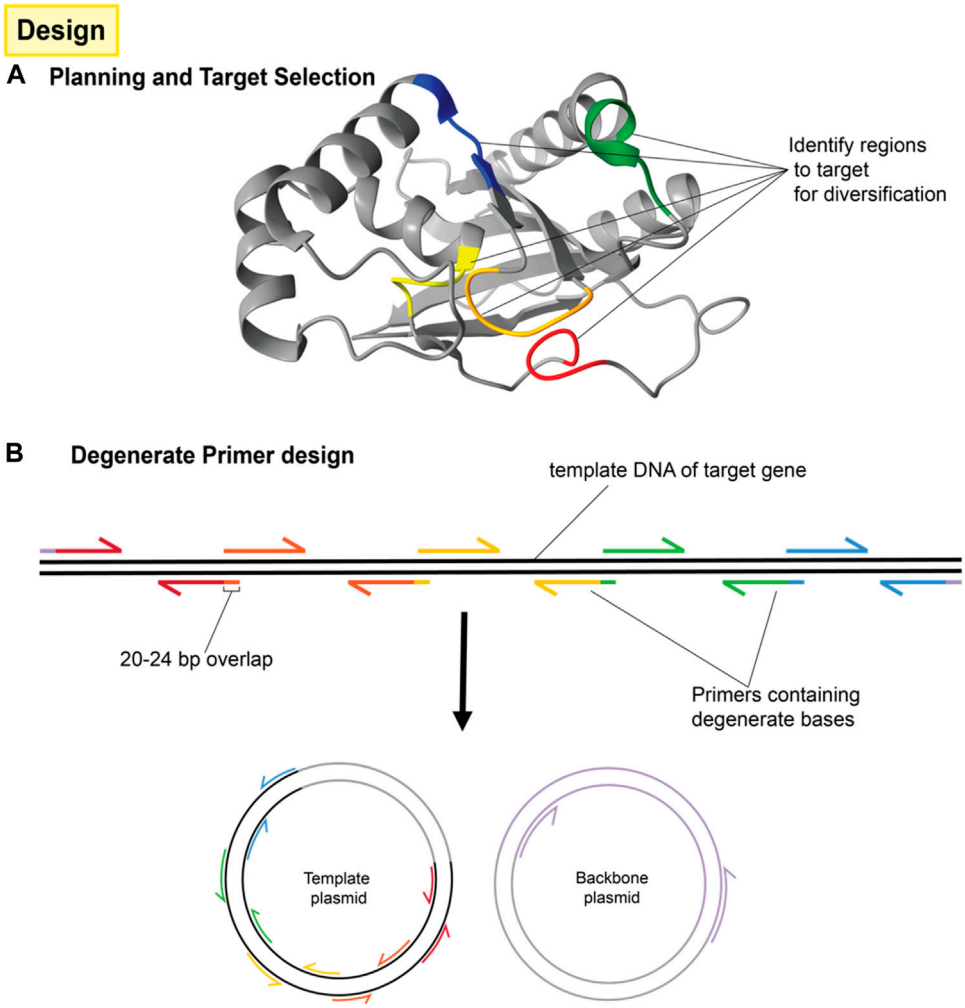


FIGURE 3
Design of protein mutation library and primers. **(A)** A number of residues are targeted for mutagenesis around the area of interest (typically a binding site or active site) but although those residues may be close to each other in the three dimensional structure, they may be far apart in the primary sequence. **(B)** The protein sequence is then separated into several fragments with the mutation sites located at one end of each fragment. This usually requires $n + 1$ fragments for n mutations unless one of the mutations is very near either end of the protein sequence. Each fragment will be generated by PCR reactions consisting of the template to be mutated and a pair of primers. One of those primers (the mutation primer) includes degenerate bases in the middle to effect mutation. The mutation primer must also include sufficient non-mutated bases on the 3′ end for binding and must include a 5′ overlap to the adjacent fragment. The other primer (the helper primer) does not need to contain degenerate bases and simply allows for the amplification of that fragment. Fragments are then combined by overlap extension PCR. Terminal primers include overlaps for cloning into the intended vector.

found with a higher frequency in nature than expected by chance, while a negative score indicates that this alignment is less frequent than expected. Practically, a positive score represents a statistically conservative substitution, while a negative score represents a non-conservative one. In the case of the BLOSUM62 matrix, the scores shown were determined based on sequences with an identity of 62% or less – making it useful for generating variants that are dissimilar to the starting sequence, but not entirely divergent (Eddy, 2004). Statistically conserved mutations form the basis of neutral drift mutations in an evolutionary trajectory, resulting in gain of functions (as in paralogs or orthologs).

The method described here lends itself well to generating large protein libraries, even based on limited initial information. Due to its capacity for screening the combinatorial effects of a broad range of mutations at once, it is possible to target several sites in a single protein for mutagenesis without significantly increasing the workload involved. This can be especially useful when constructing a new biosensor, where one may wish to target its capacity for ligand-binding, DNA-binding and multimerization all at once (Figure 3). This approach to protein engineering for biosensor development was successfully applied in the development of a protocatechuate biosensor in *P. putida* using an IclR transcription factor PcaU (Figure 2B) (Jha et al., 2018).

1.3 Applications of the method

The methodology outlined in this paper is broadly applicable to a range of objectives. Within the limitations outlined below (discussed in Section 1.4), we envision that this high-throughput

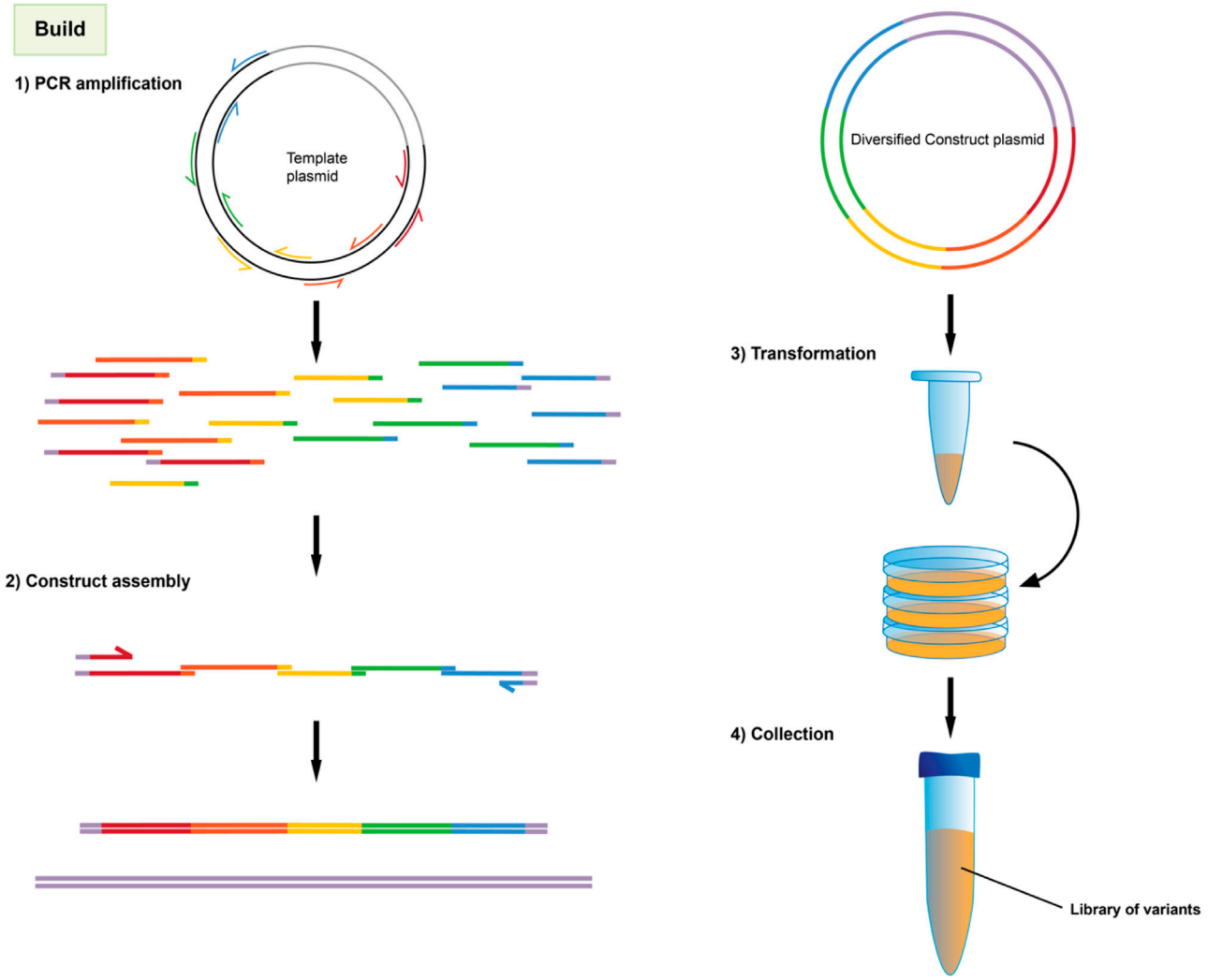


FIGURE 4
Overview of library construction method.

approach using large libraries of diverse genetic variants, can be successfully employed in any project requiring random, or semi-rational design of pathways, proteins, and promoter variants. We have successfully applied this method to optimize promoter activity (Bentley et al., 2020; Pardo et al., 2020; Shin et al 2022), alter ligand binding (Jha et al., 2016; Shin et al 2022), reduce enzyme inhibition (Jha et al., 2019), construct biosensors (Jha et al., 2014; Jha et al., 2015) increase enzyme efficiency (Jha and Strauss, 2020), and improve thermostability and expression of an enzyme (Harrington et al., 2017), while taking advantage of the high throughput efficiency of flow cytometry.

1.4 Comparisons and limitations

1.4.1 This method

The method described here introduces a variety of mutations at each mutation site by performing PCR-based, site directed mutagenesis, with primers containing degenerate bases at the desired mutation site. Multiple residues of the protein or positions in the promoter can be targeted simultaneously, with each targeted mutation region produced by a separate PCR reaction. Diversity is then further enhanced by the combinatorial assembly of those fragments by overlap extension PCR (Bryksin et al., 2013). The resulting gene or promoter library is then cloned into a vector by Gibson assembly (Gibson et al., 2009). While none of these three components (site directed mutation with degenerate primers, overlap extension PCR and Gibson assembly) are novel in isolation, their combination is rare within the mutation library field, despite offering superior flexibility, control of mutation bias, and ease of use when compared to the alternatives below.

1.4.2 Alternatives to rational library design

The primary function of the mutation library is to create genetic diversity that can then be screened, tested, or even fed into some biological selection process. Genetic diversity can also be achieved by random mutagenesis promoted by chemical mutagens or radiation

(Zhang et al., 2018). Given enough generations, mutations can also be accrued naturally, especially during adaptive laboratory evolution (ALE) (Zheng et al., 2021). The advantage of these alternative methods is that they do not require much up-front design work or cloning, although these techniques can be enhanced by selection methods, such as growth coupling (Godara and Kao, 2020) or biosensors with antibiotic resistance response (Yin et al., 2022).

The disadvantage of using these alternate methods lies in their more random nature. Without designing a bias towards potentially useful mutations, it is expected that a much larger percentage of the mutations will be deleterious, which requires screening or selecting from a larger sampling of the population. It is important to note, however, that ALE, random mutagenesis and mutation libraries are not mutually exclusive methods, but can complement each other. For instance, random mutagenesis, semi-rational libraries and even fully rational mutation libraries have been used to provide the genetic diversity for ALE to act upon (Arora et al., 2020).

1.4.3 Alternative methods for targeted mutagenesis and library construction

There is a plethora of methods currently available to generate both random and targeted mutagenesis. Targeted mutagenesis approaches, such as site directed mutagenesis (SDM) and site-saturation mutagenesis (SSM) have proven to be extremely useful, however, the relatively low library size that can be achieved through these methods has restricted their application in directed evolution approaches (Sayous et al., 2020). Most recent technologies, such as sequence saturation mutagenesis (SeSaM), where a universal base is inserted along the target sequence, randomizing it at every single position (Wong et al., 2004), or casting error-prone PCR (cepPCR), where target DNA is fragmented and amplified using error-prone PCR (Yang et al., 2017), have increased achievable library sizes and mutational coverage. Yet, all of these methods required labor-intensive steps of cloning and transformation. The unprecedented drop in cost of DNA synthesis has allowed the generation of DNA libraries by high-throughput oligo synthesis (Kosuri and Church, 2014; Rocklin et al., 2017), enabling a complete saturation of small proteins. Still, the high price, compared to other techniques, makes this technique appropriate for only some specific applications.

After mutations have been generated using one of the above methods, a variety of cloning methods can be used for insertion into a replicating plasmid or genome. Traditional restriction/ligation methods have largely given way to PCR-based, recombination-based and CRISPR-based methods. The method described here uses Gibson assembly, which is versatile and familiar to many synthetic biology labs, but the protocol described herein could be easily adapted to use via megaprimer methods for insertion such as MEGAWHOP. The MEGAWHOP method traditionally consists of two PCR steps. The first step uses error-prone PCR for the generation of a set of megaprimers with random mutations in the target gene. The second step of PCR uses the megaprimers and the original plasmid as the template, resulting in a large random mutagenesis library (Miyazaki and Takenouchi, 2002; Miyazaki and Voigt, 2011). Compared to Gibson assembly, megaprimer methods have the advantage of requiring fewer enzymes and do not need a linearized backbone for insertion. Disadvantages include a higher incidence of mutations in the backbone as it is replicated by PCR. More advanced versions of the megaprimer method, such as QuickStep cloning (Jajesniak and Wong, 2015), have several advantages over earlier iterations, including exponential amplification of the whole plasmid and lower chances of self-annealing of the megaprimer at the 3′ ends.

Regarding chromosomally-targeted mutagenesis, homologous recombination (Recombineering) (Thomason et al., 2014) is the most commonly used technique. Recombineering requires specific single-stranded DNA annealing proteins that are highly specific and whose efficacy varies among different bacterial species, in addition to *in vitro* methods to generate sequence diversity.

CRISPR-based methods may also be used to mutate the genome directly. These methods exploit the sequence-specific mode of action of CRISPR, generally in combination with Cas9. For instance, CRISPR-enabled trackable genome engineering (CREATE) (Garst et al., 2017) is a method that makes use of large-scale oligonucleotide synthesis to generate a pool of 10^4–10 (Salis et al., 2009) barcoded oligonucleotides, and then uses CRISPR to achieve mutagenesis within the genome. Additionally an *in vitro* CRISPR/Cas9-mediated mutagenic (ICM) system for construction of designer mutants in a PCR-free approach has been reported (She et al., 2018). In this method, CRISPR/Cas9 is used to cleave plasmid DNA at a target site, followed by T5 exonuclease digestion and annealing of primers containing the intended mutations. In both cases, CRISPR/Cas9 is used to cut the DNA strands, while the genetic diversity is achieved by synthetic DNA oligonucleotides containing the desired mutations. Both CRISPR and gene synthesis methods can eliminate bias when creating DNA libraries, however, PCR methods, such as the one presented here, remain the most widely applied, due to their low cost and ease of use (Table 1). Furthermore, the method described here allows for control of bias mutations, since the design of degenerate primers will allow saturation or partial saturation mutagenesis. This method can be used in non-model organisms, helps achieve both combinatorial and targeted mutagenesis, can be performed using relatively simple and widely used molecular biology techniques, and can also be employed in combination with other mutagenesis methods described in Table 1.

1.4.4 Limitations of this method

The method we present here is capable of quickly generating large libraries of genetic variants and quickly screening them for desired phenotypes. However, the construction of the library relies on degenerate primers. These are produced in such a way as to control the ratio of nucleobases at a given location, but the different sequences may show differences in annealing during PCR resulting in a degree of bias. Ensuring sufficient 3′ complementarity after the mutation site mitigates this effect, but the exact ratio in the final library is not always known. Thus, this protocol includes a verification step of either sequencing a random selection of clones isolated from the library (see Figure 6A), or sequencing DNA extracted from the whole library (see Figure 6B), for quality assurance. This provides reasonable certainty that the desired mutations were achieved, and that the library contains a sufficiently broad range of variants to be of interest for screening.

This method is additionally limited by availability of knowledge about the targeted protein or promoter (see Section 1.2), as well as by

TABLE 1 Comparison of methods used for generation of mutation libraries.

Name (Acronym)	Description	Advantages	Disadvantages
Site-directed mutagenesis (SDM) (Sayous et al., 2020)	A specific mutation is introduced in a gene by primers. One or two adjacent amino acids can be mutated per primer	Focused and very specific. Ideal for targeting single bases for mutation. Useful for the creation of small libraries (8.5 h–2 days for 10–100 variants) (Bachman, 2013). High accuracy (>80%) if using high-fidelity DNA polymerase (Alejaldre et al., 2021)	Each target mutation requires an individual mutagenesis reaction, thus cost and time limit the achievable library size. Only useful if the target protein is already very well characterized. Efficiency relies on the availability of suitable cloning systems (Alejaldre et al., 2021)
Site-saturation mutagenesis (SSM) (Sayous et al., 2020)	One amino acid is mutated to the other 19 amino acids by whole-plasmid PCR with premers consisting of degenerate codons (Kretz et al., 2004), achieving all possible mutations for a single position	Larger genetic diversity than SDM (Siloto and Weselake, 2012). Reduces mutational bias	Workload similar to SDM, requires individual mutagenesis reaction and transformation for each target residue, not combinatorial
Sequence saturation mutagenesis (SeSaM) (Wong et al., 2004)	DNA fragments with a random spread of sizes are generated and elongated at the 3′ end using a universal base. The universal base is then randomly replaced with standard nucleotides during PCR, generating a randomized sequence	Large diversity of mutants, can theoretically achieve full randomization of a desired sequence	Moderately labor-intensive (2–3 days to generate a library), involved multi-step PCRs, universal base introduces mutation bias, interval of mutations depends on fragmentation method used, not combinatorial, introduces stop codon at a frequency of 0.03
Casting error-prone PCR (cepPCR) (Yang et al., 2017)	Target DNA is amplified into 100-200 bp fragments. Each fragment is subjected to high-rate error-prone PCR. Mutant fragments are then used as megaprimers (see MEGAWHOP) to generate a library of mutants for each fragment	Reduces redundancy of mutations, achieves higher mutation rate than traditional error prone PCR.	Moderately labor-intensive (1–2 days to generate libraries), mutation bias, only achieves a subset (<40%) of potential beneficial mutations. The multiple resulting libraries may need to be screened individually, limiting this method to shorter sequences
Megaprimer PCR of whole plasmid (MEGAWHOP) (Miyazaki, 2011)	A megaprimer carrying desired mutations is hybridized to the template DNA for whole plasmid PCR. The original template is then degraded by *DpnI* digestion	Produces variants with a combination of mutations. Can be combined with cepPCR or degenerate nucleotides, no need for ligation or assembly	Synthesizing megaprimers may carry high costs, In-house generated megaprimers have the same shortcomings as the method used to generate them (i.e., epPCR or cepPCR). Higher chance of introducing mutations in the backbone
Library DNA synthesis (Kosuri and Church, 2014; Rocklin et al., 2017)	DNA libraries are created by high-throughput oligo synthesis to user specifications	Eliminates mutation bias, large library size ($>10^6$) and custom-fit. Reduced workload for user	High cost ($0.2-0.5 per bp for large genes). Potentially longer lead times to receive library
Homologous recombination (Recombineering) (Thomason et al., 2014)	A previously made DNA fragment is inserted in the genome by homologous recombination	Mutants can be introduced in the genomic DNA, improving stability and eliminating the need for reporters	Needs highly specific DNA annealing proteins, and optimization
CRISPR-enabled trackable genome engineering (CREATE) (Garst et al., 2017)	CRISPR/Cas9 is used to cut genomic DNA, and barcoded oligonucleotides are inserted in the desired position	Genome editing at multiple loci, mutations can be tracked. Allows >50,000 genome wide mutations (Tarasava et al., 2018). Used in *E. coli* and *S. cerevisiae*. Off-target mutations are expected to be low, no mutation bias, relatively easy to perform	Requires potentially expensive barcoded oligo synthesis (ex: ~$5000 for a 10^5-variant library of 50bp oligos (Garst et al., 2017)). Low efficiency (>70%), requires development of Cas system, has not been demonstrated in mutagenesis of single, specific enzyme
In vitro CRISPR/Cas9-mediated mutagenic (ICM) (She et al., 2018)	CRISPR/Cas9 is used to cut plasmid DNA, and then mutation carrying oligonucleotides are inserted in the desired position	Off-target mutations are expected to be low, no mutation bias, no need for restriction enzyme digestion, PCR, restriction sites or plasmid size limit	As with CREATE, high cost due to the need for high-throughput oligo synthesis
This method	Combines overlap extension PCR and saturation or partial saturation mutagenesis with degenerate primers	Allows for rapid in-house creation of large, combinatorial variant libraries ($\geq 10^6$), bias mutations can be controlled, low cost (e.g., only a dozen 18–50 bp primers for 5 mutation sites spread over a gene or fewer primers if some mutations are near enough to each other such that they may be covered by a single primer)	Moderately labor intensive (2–3 days for library construction)

the expertise of prospective user regarding FACS. Additionally, if looking to build a biosensor/binder for a molecule that is readily metabolized or exported, then it may be challenging to achieve an intracellular concentration sufficient for screening by FACS, and metabolic engineering to disable one or more metabolic pathways may be required.

1.5 Expertise needed

In order to successfully apply this protocol, the prospective user will need experience with basic molecular biology methodologies, including: polymerase chain reaction, agarose gel electrophoresis, gel extraction, bacterial transformation, and bacterial culture, as well as a fundamental understanding of molar ratios and calculations. Experience using FACS is required. Additionally, familiarity with primer design and gene editing software is necessary. Sequences for the targeted gene or promoter need to be known or obtainable.

When targeting a gene for diversification, it is helpful to have an understanding of the protein encoded therein, and its structure-function relationship. Familiarity with Rosetta (Rohl et al., 2004) or AlphaFold (Jumper et al., 2021) and its protein folding and ligand docking functionality is beneficial for rational protein diversification approaches. Similarly, for promoter engineering, an understanding of promotor features and transcription factor mechanisms assists with targeting specific regions for mutagenesis (Browning and Busby, 2004; Saecker et al., 2011).

2 Materials

2.1 Reagents

- Oligonucleotides can be purchased from Eurofins or other vendors. Oligonucleotides are dissolved in ultrapure water to a concentration of 50 μM
- UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, cat. no. 10977015)
- Deoxynucleotide (dNTP) solution mix (New England Biolabs, cat. no. N0447L)
- High-fidelity DNA polymerase (e.g., New England Biolabs, Phusion DNA polymerase, cat. no. M0531L, Q5 DNA polymerase Q5® High-Fidelity 2X Master Mix, cat. no. M0492L)
- Dimethyl sulfoxide (DMSO) Fisher BioReagents, cat. no. BP231-100
- Agarose (e.g., Invitrogen, UltraPure™ Agarose, cat. no. 16500100)
- Gel Loading Dye, Purple (6X), no SDS, cat. no. B7025S
- DNA ladder (e.g., New England Biolabs, 1 kb DNA Ladder, cat. no. N3232L, 100 bp DNA Ladder, cat. no. N3231L)
- GelRed® Nucleic Acid Stain 10,000X Water (Sigma, cat. no. **SCT123**)
- TAE Buffer (Tris-acetate-EDTA) (50X) (Thermo Fisher Scientific, cat. no. B49)
- QIAquick Gel Extraction Kit (QIAGEN, cat. no. 28706X4)
- QIAquick PCR Purification Kit (QIAGEN, cat. no. 28104)
- MinElute Reaction Cleanup Kit (QIAGEN, cat. no. 28204)
- Restriction enzymes and 10X reaction buffer
- T4 DNA ligase (New England Biolabs, cat. no. M0202L)
- Antarctic Phosphatase (New England Biolabs, cat. no. M0289S)
- High-efficiency bacterial competent cells (e.g., Thermo Fisher Scientific, MAX Efficiency™ DH5α, cat. no. 18258012)
- Assembled plasmid DNA library, user supplied.
- SOC Outgrowth Medium (New England Biolabs, cat. no. B9020S,
- Bacterial growth and selection medium (liquid and agar) e.g., Luria Broth Base (Thermo Fisher Scientific, Miller's LB Broth Base, cat. no. 12795027
- Antibiotics (e.g., kanamycin sulfate, Thermo Scientific, cat. no. 11815024)
- Phosphate buffered saline (PBS) (G-Biosciences, cat. no. 786-027)

2.2 Equipment

- Incubators at appropriate temperature and agitation
- Thermocycler (e.g., Applied Biosystems 2720 Thermal Cycler)
- Gel electrophoresis system (e.g., Thermo Fisher Scientific, Owl™ EasyCast™ B1A Mini Gel Electrophoresis Systems, cat. no. B1A-BP)
- ChemiDoc Imaging System (Bio-Rad)
- Cell scrapers
- Tube rotator
- Flow cytometer (FACSAria III flow cytometer) capable of cell sorting based on fluorescence

2.3 Standard laboratory consumables

- PCR tubes
- 1.5 and 2 mL Eppendorf tubes
- 14 mL culture tubes
- Petri dishes

2.4 Software

- SnapGene or similar
- Protein modeling software such as ChimeraX or PyMOL
- Rosetta or AlphaFold

3 Procedure

3.1 Overview

This protocol can be divided into three main components: Library Design, Library Construction, and Library Screening analogous to the Design, Build, Test framework of engineering principles for synthetic biology (Peccoud, 2016; Opgenorth et al., 2019). The three main components can be further divided into individual steps as described in Figure 1. The principals associated with Library Design were discussed in Sections 1.1, 1.2 above. The protocols for Library Construction and Library Screening are detailed below.

In total, this method will take approximately 3 weeks from initial library design to isolating and characterizing individual clones, for a reasonable library size of 10^5–10^6, with the workload for each day itemized below. Depending on the library size, a smaller or larger workload can be expected for smaller or larger libraries respectively.

Day 1: PCR round 1 to generate fragments. Agarose Gel electrophoresis and gel extraction.

Day 2: PCR round 2 to assemble, Agarose Gel Electrophoresis and gel extraction,

Day 3: Transformation (include main plates and transformant estimation plates).

Day 4: Plate scraping (and colony counting), direct use or glycerol stock, extractions for sequencing, inoculation of liquid culture.

Day 5: Re-inoculation and induction.

Day 6: First round of analysis by flow cytometry and sorting, followed by outgrowth (NOTE: In order to collect rare clones, in the first round of sorting it is recommended to collect the top 5% of the library).

Day 6: Glycerol stocks, liquid culture of round 1 populations.

Day 7: Re-inoculation of round 1 populations, induction.

Day 8: Second round of analysis and sorting, outgrowth (NOTE: A negative sorting is recommended to eliminate constitutive performers, especially observed when engineering regulatory proteins and promoters).

Day 9: Glycerol stocks and liquid culture of round 2 populations.

Day 10: Re-inoculation and induction.

Day 11: Third Round of screening and sorting, outgrowth (NOTE: Increased stringency in sorting, i.e., collecting only the top 1%–2% of population is recommended).

Day 12: Glycerol stocks and liquid culture of round 3 populations.

Day 13: Re-inoculation and induction.

Day 14: Fourth round of screening and sorting, plating with appropriate antibiotic selection (NOTE: Stringently collecting only the top 1% of performers is recommended).

Day 15: Picking 24-48 colonies of individual clones.

Day 16–19: Test individual clones at different conditions in order to characterize properties.

3.2 Primer design and construction

The following instructions detail the production of a hypothetical mutation library, wherein two distant regions of a promoter or protein are diversified. Mutation sites can range in size from a single base pair up to any stretch of the sequence reasonably covered by a single PCR primer after factoring in the 3′ overlap needed for the initial PCR and the 5′ overlap needed for overlap extension PCR described below.

3.2.1 PCR based mutagenesis: (1 day)

1. Primers are designed according to standard site-directed mutagenesis and overlap extension PCR strategies (Ho et al., 1989; Heckman and Pease, 2007). For our example of two mutation sites, a total of six primers are needed; one degenerate primer at each mutation site with 20 base pair overlaps that extend into adjacent fragments, one non-variable primer for each mutation site to serve as the reverse primer for that fragment and finally, two primers to flank the entire region. This divides the promoter into three fragments separated by mutation sites. Figure 3B shows the same concept, but for four mutation sites. NOTE: It is sometimes necessary to use more degenerate primers at each mutations site. For instance, if the library is designed to include three possible codons at a given amino acid position; GCA, CAG, and GAA (for alanine, glutamine and glutamic acid, respectively), then it would *not* be appropriate to use a single primer containing the degenerate bases SMR (See Supplementary Table S2), because these could also combine to code for proline. Instead, two different primers containing GMA or CAG could be used and mixed 2:1 in the PCR reaction for all three amino acids to be equally represented.
2. PCR is performed for each fragment separately. For primers with degenerate bases, annealing temperatures should be lowered to accommodate the mutation variant with the least stable hybridization to the template. Follow suggested thermocycler settings for whichever high-fidelity polymerase is used.
3. The entire PCR product is then run on an agarose gel and the correct sized bands are excised.
4. Extract DNA from the gel excisions using commercially available gel extraction kits (e.g., QiaQuick gel extraction kit)

3.2.2 Construct assembly (1 day)

5. PCR fragments are then combined via overlap extension PCR. In the first stage of overlap PCR, the fragments (at equimolar ratio) and PCR reagents/enzymes are allowed to react for 8 cycles with an extension time sufficient to copy the largest fragment. Then, primers flanking the entire promoter region are introduced and 25 more cycles are performed with an extension time sufficient for the entire region. Subsequent purification of the PCR product with commercial kits assists the next step (i.e., QiaGen PCR Cleanup Kit).
6. The mutated variant library is then assembled into a linearized vector using Gibson assembly (Gibson et al., 2009; Thomas et al., 2015) or restriction digestion/ligation. Restriction sites may be introduced into the primers used to flank the promoter in step 5.

3.2.3 Transformation (1 day)

7. The assembled plasmid library is transformed into a suitable competent bacterial strain, Transformation may be performed by heat shock or electroporation. Depending on the known or expected transformation efficiency, care must be taken to perform sufficient transformations to achieve appropriate coverage of the library. Commonly one would aim to obtain at least 4-fold coverage e.g., generating ≥1 million transformants for a library with a theoretical diversity of 250,000. This is to ensure that the maximum number of variants is represented in the bacteria.

After transformation and recovery, it is advisable to plate approximately 20 μL of the recovered bacteria on an agar plate containing the appropriate medium and selective antibiotic. This is to estimate transformation efficiency and therefore the final degree of coverage of the library that was achieved (i.e., quantification plate). The remaining recovered bacteria are gently spun down (4,000 rpm, 4–5 min), and the majority of the supernatant recovery medium is removed to allow for plating of the entire volume of transformants. Once plated, the transformants are incubated overnight at a suitable temperature to form colonies.

3.2.4 Collection and stocks (1 day)

8. Using the quantification plate from step 7, the number of transformants achieved is estimated. If the desired coverage is achieved, the whole library can be pooled by adding a small amount of liquid medium to the plates (typically 1 mL), and then gently scraping the colonies to collect, using a cell scraper. Repeat the addition of liquid media, scraping and then pool the resulting cell suspensions in a polypropylene tube (15–50 mL, depending on resulting volume), sealed tightly, and then rotated for at least 1h to ensure proper mixing of the collected library.
9. Glycerol stocks of the collected library are prepared by adding glycerol to aliquots of the library to a final concentration of 20% glycerol, taking note of final OD of the resulting stock. These stocks are suitable for long-term storage in ultracold freezers (−80°C). Note that when reviving culture from stocks it is essential to use sufficient inoculum to achieve full coverage of the library i.e., inoculate fresh culture with a number of cells at least 10-fold greater than the total number of variants in the library.

3.2.5 Verification (2–5 days)

10. In order to verify the integrity and diversity of the library, plasmid DNA is isolated from individual clones (e.g., picked from the quantification plate in step 7 and grown up overnight) or from a small volume of the collected library, using any desired plasmid DNA extraction method (i.e., QiaGen Miniprep Kit). The isolated DNA, along with an appropriate primer, is sent for sequencing by one's preferred provider (in-house, Twist, Eurofins, etc.). Depending on usual shipping and processing times, it may take several days for the sequencing data to become available. See Figures 6A, B for an example of expected results.

3.3 Library Screening

3.3.1 Two rounds of positive selection (3–5 days)

11. The library should be screened using media and other growth conditions mirroring the application and desired effect of the mutations. This method assumes testing of cells at mid-log growth phase. Prepare an overnight culture in 3 mL of selective liquid media, using either the scraped cell suspension from step 8 (if available) or the glycerol stock of the library. Use a sufficiently large volume of inoculum to achieve library coverage (see Step 9).
12. Use the overnight culture to inoculate a fresh 3 mL culture to an initial OD that is about 1/10th of the strain's stationary phase OD in that media. Incubate the cells with frequent OD monitoring until they reach approximately 50% of stationary OD (mid-log phase). If induction is required, induce at mid-log phase and allow more time for the induced process to proceed.
13. Dilute a small portion of the cells in 1X phosphate buffered saline to achieve a cell density of approximately 10^7cells/mL for flow cytometry. Required dilution may vary depending on the requirement of the instrument used, since efficient sorting requires an event rate well below the recommended maximum event rate for any given flow cytometer.

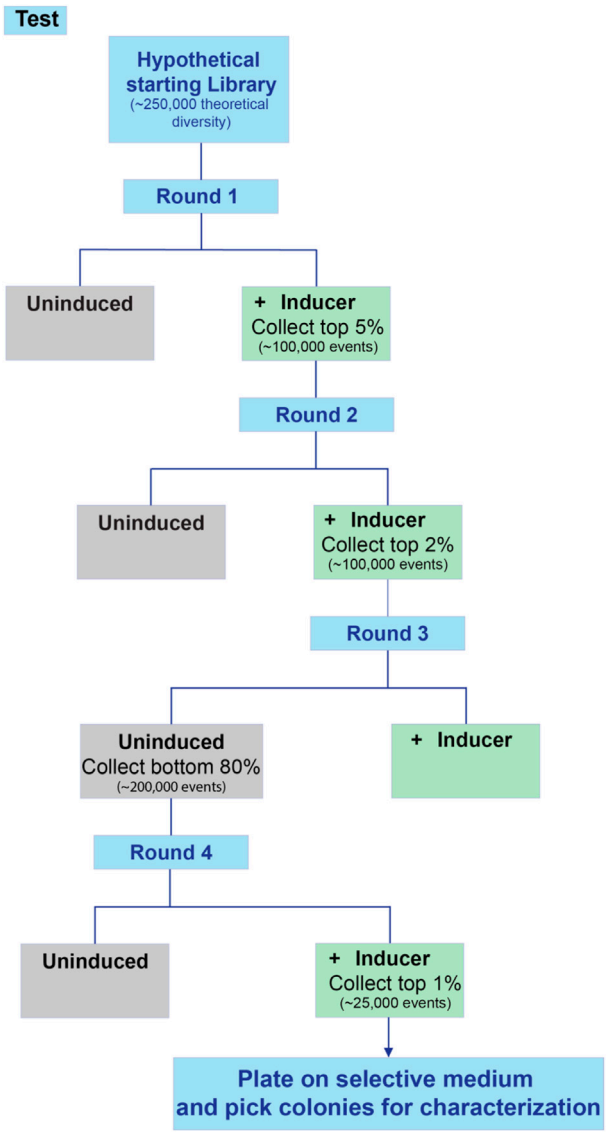


FIGURE 5
Workflow for screening an inducible promoter library. When screening inducible promoters, it is advisable to include several rounds of positive selection (high expression when induced) as well as at least one round of negative selection (low background expression when uninduced). This example includes two rounds of positive selection, one round of negative selection, and a final round of positive selection.

14. Select the top five percent best performing cells based on biosensor response and sort into fresh (selective) media. Grow the sorted cells overnight to recover.
15. Repeat steps 12-14, this time collecting the top two percent only.
16. If induction was used, or the library is for a new biosensor or promoter to be optimized, it may be necessary to screen for negative fluorescence in the uninduced state i.e., select a non-fluorescent subset of the uninduced population. This will eliminate constitutively active variants, facilitating isolation of true inducible variants. If deemed necessary, repeat steps 12-

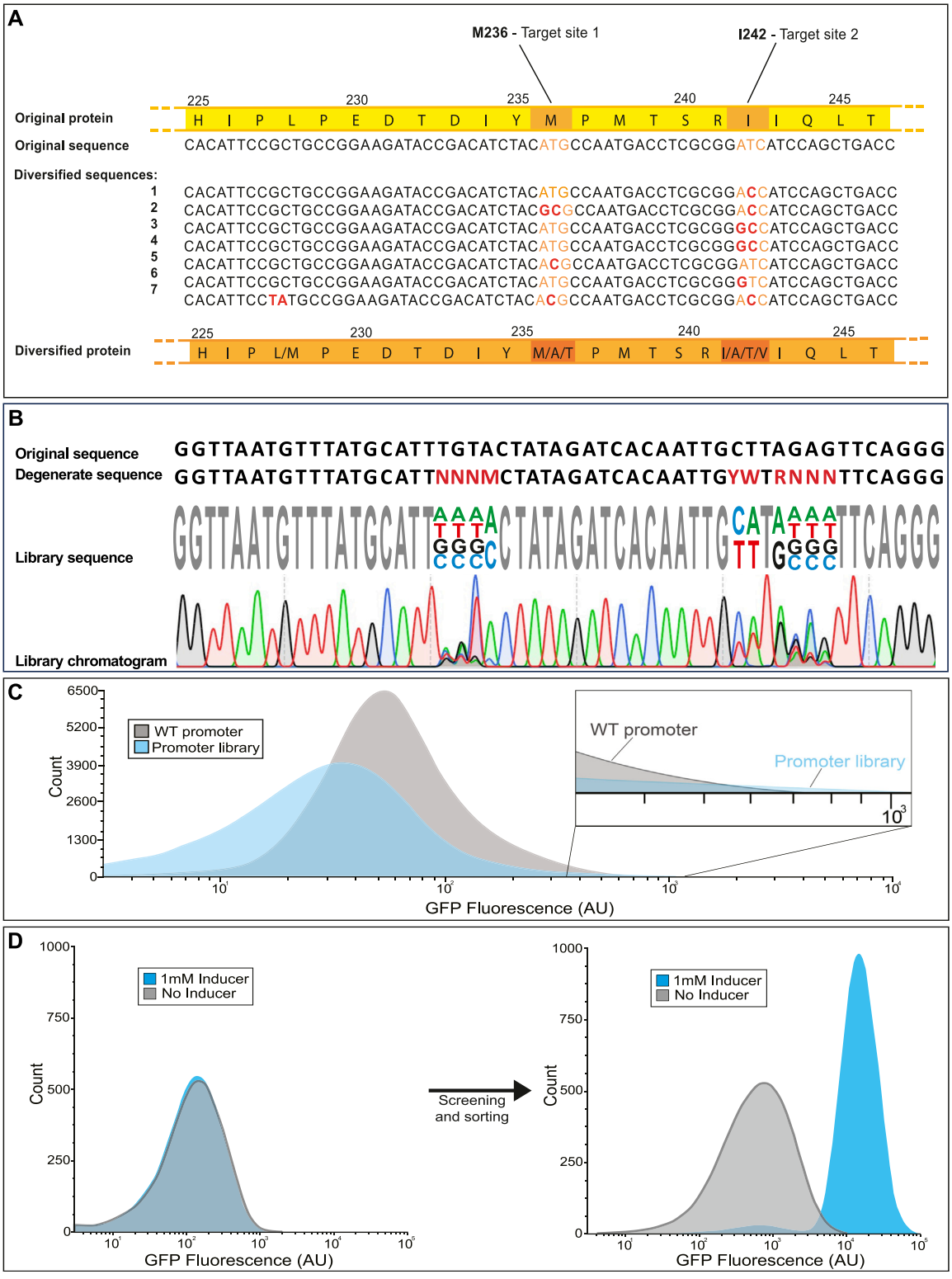


FIGURE 6
Expected results for library validation in various applications. The library is first sequenced to ensure that each mutation site exhibits the expected variability. This can be performed in two different ways; **(A)** by sequencing several isolates and aligning them for comparison or **(B)** by sequencing the mixed population and observing overlapping peaks in the chromatogram. **(C)** Ideally the library exhibits a broader distribution of phenotype then that of the parent strain. It is acceptable if many of the mutations are deleterious as long as the rare strains with improved performance can be selected from the library (See inset). **(D)** After selection, high performing mutants can be compared to the wild type by cytometry. An example for isolating a novel biosensor by diversifying a wild type transcription factor is shown. The wild type transcription factor CatM (right) shows no response to the added inducer, while the variant isolated after several rounds of screening (left) displays a >40-fold increase in fluorescence in response to the inducer (Bentley et al., 2020).

14 without induction and collect the bottom 80% in terms of fluorescence.

17. Alternate between induced selection and uninduced selection until the desired phenotype is reached or no further improvement is observed between rounds of sorting. An example screening flowchart for an inducible process is shown in Figure 5. Note that sorts with a wider selection window in terms of percentage can include more cells collected, but that sorts with narrow windows have fewer cells collected in the interest of time.
18. Plate cells from the last sort onto selective media and pick individual colonies for characterization.

4 Expected results

During construction of the library, there will be multiple tests performed at intermediate phases to ensure the components are being generated and assembled correctly. During the initial PCR steps, amplification products are observed on agarose gel and can be checked for the correct size. After the gene is inserted into a vector by Gibson assembly and transduced into a host, a subset of the transformation culture is diluted and used for counting plates (quantification plate) which have the additional benefit of providing isolated colonies from which to test the diversity of the library at each mutation site prior to employing high throughput screening or selection. Before screening, a check may be included to test for mutation diversity (Figure 6). One example, shown in Figure 6A involves selecting several random isolates from a library which are sequenced (Sanger method) at the mutation sites. Given the degenerate bases provided, there are three possible amino acids (including the wild type) for the inducer binding residue M236 and four possible amino acids for the dimer interface I242. For each mutation site, every possible amino acid substitution was observed in the seven isolates. Alternatively, the library may be sequenced directly, without isolating individual strains. In this approach, degenerate bases should produce multiple peaks of different fluorophores on the sequencing chromatogram depending on the rate of base-pair substitutions (Figure 6B).

Based on the goals set forth for the library design, the library is expected to exhibit a broad distribution of phenotype in addition to genotype. In the case of libraries whose performance can be tested and screened by fluorescence, the phenotypic range can be measured on a flow cytometer. In the example histogram provided (Figure 6C), the majority of mutation combinations for a promoter library are deleterious, causing the average fluorescence to go down. The strength of the library lies in the rare mutations with gain of function or improved performance, evident in the high fluorescence tail of the library population which extends slightly beyond the high fluorescence region of the wild type population. By selecting cells from those regions of the library histogram that perform better than the wild type, it is possible to isolate variants with markedly improved function. One example is a cis,cis-muconanic acid biosensor isolated from a biosensor library (Figure 6D). While the wild-type transcription factor (Figure 6D left) showed no visible response to the inducer, the variant isolated from the library (Figure 6D right) responded remarkably well to induction, with only a minor shift in background fluorescence compared to the wild-type.

5 Conclusion

Current methods designed to identify gain-of-function mutations for proteins or promoters in microbes are limited by workload, time, and costs. The methodology described here eliminates the bottleneck of testing individual variants of genes and promoters by presenting a protocol for assaying the combinatorial effect of several mutations at once, drastically reducing the time and money required to obtain genetic variants with desirable qualities. The efficacy of the method described here is illustrated by several of our original research papers (Jha et al., 2014; Jha et al., 2016; Jha et al., 2018; Shin et al., 2022), which have successfully employed this approach. The versatility and usefulness of this high-throughput screening method applied to large genetic libraries is therefore evidenced in published literature. While the individual components of this workflow (overlap extension PCR, degenerate primers, FACS, etc.) may seem commonplace, we have not encountered any other methodology that combines them in the fashion outlined in this paper.

Due to the high-throughput nature of the method detailed here, the number of genetic variants that can be screened for desired gain-of-function behavior is not limited by the time or manpower available. Instead, millions of variants can conveniently be screened for desirable phenotypes in a single sample tube. Variants that perform well are isolated by FACS and screened further in subsequent rounds, while poor performers are easily discarded.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

All authors developed and tested the method for various protein engineering, biosensor development and metabolic engineering efforts in several microbial strains, including *E. coli*, *P. putida* and *A. baylyi*. HH, E-KT, and RJ conceptualized the manuscript. RJ and TD acquired funding. All authors contributed to the article and approved the submitted version.

Funding

The work was authored under Triad National Security, LLC ("Triad") Contract No. 89233218CNA000001 with the U.S. Department of Energy. The work was supported through the Agile BioFoundry (ABF), which is funded by the US Department of Energy, Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office (BETO) under contract

NL0032182, and through Bio-Optimized Technologies to keep Thermoplastics out of Landfills and the Environment (BOTTLE) Consortium, supported by DOE- EERE's Bioenergy Technologies Office and Advanced Materials and Manufacturing Technologies Office under Contract nos. NL0035994 and NL0037843, respectively. The work was also supported by the US Department of Energy, Office of Science Biological and Environmental Research program under contract number FWP LANLF32A.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1202388/full#supplementary-material

References

Alberghina, L., and Lotti, M. (2005). "Protein engineering in basic and applied Biotechnology: A review," in *Protein engineering and industrial Biotechnology* (Harwood Academic Publishers), 1–21.

Alejaldre, L., Pelletier, J. N., and Quaglia, D. (2021). Methods for enzyme library creation: Which one will you choose?: A guide for novices and experts to introduce genetic diversity. *BioEssays* 43, 2100052. doi:10.1002/bies.202100052

Arora, N., Yen, H.-W., and Philippidis, G. P. (2020). Harnessing the power of mutagenesis and adaptive laboratory evolution for high lipid production by oleaginous microalgae and yeasts. *Sustainability* 12, 5125. doi:10.3390/su12125125

Bachman, J. (2013). "Site-directed mutagenesis," in *Methods in enzymology* (Elsevier), 529, 241–248.

Bakke, I., Berg, L., Aune, T. E. V., Brautaset, T., Sletta, H., Tøndervik, A., et al. (2009). Random mutagenesis of the Pm promoter as a powerful strategy for improvement of recombinant-gene expression. *Appl. Environ. Microbiol.* 75, 2002–2011. doi:10.1128/aem.02315-08

Bentley, G. J., Narayanan, N., Jha, R. K., Salvachúa, D., Elmore, J. R., Peabody, G. L., et al. (2020). Engineering glucose metabolism for enhanced muconic acid production in *Pseudomonas putida* KT2440. *Metab. Eng.* 59, 64–75. doi:10.1016/j.ymben.2020.01.001

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235

Blazeck, J., and Alper, H. S. (2013). Promoter engineering: Recent advances in controlling transcription at the most fundamental level. *Biotechnol. J.* 8, 46–58. doi:10.1002/biot.201200120

Boldrin, F., Degiacomi, G., Serafini, A., Kolly, G. S., Ventura, M., Sala, C., et al. (2017). Promoter mutagenesis for fine-tuning expression of essential genes in *Mycobacterium tuberculosis*. *Mycobacterium Tuberc. Microb. Biotechnol.* 11, 238–247. doi:10.1111/1751-7915.12875

Browning, D. F., and Busby, S. J. W. (2004). The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* 2, 57–65. doi:10.1038/nrmicro787

Bryksin, A., and Matsumura, I. (2013). "Overlap extension PCR cloning," in *Synthetic biology*. Editors K. M. Polizzi and C. Kontoravdi (Totowa, NJ: Humana Press), 31–42. doi:10.1007/978-1-62703-625-2_4

Cassiano, M. H. A., and Silva-Rocha, R. (2020). Benchmarking bacterial promoter prediction tools: Potentialities and limitations. *mSystems* 5, 00439-20–e520. doi:10.1128/msystems.00439-20

Charbonneau, M. R., Isabella, V. M., Li, N., and Kurtz, C. B. (2020). Developing a new class of engineered live bacterial therapeutics to treat human diseases. *Nat. Commun.* 11, 1738. doi:10.1038/s41467-020-15508-1

Chen, Y., Ho, J. M. L., Shis, D. L., Gupta, C., Long, J., Wagner, D. S., et al. (2018). Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nat. Commun.* 9, 64. doi:10.1038/s41467-017-02473-5

Dosoky, N. S., May-Zhang, L. S., and Davies, S. S. (2020). Engineering the gut microbiota to treat chronic diseases. *Appl. Microbiol. Biotechnol.* 104, 7657–7671. doi:10.1007/s00253-020-10771-0

Eddy, S. R. (2004). Where did the BLOSUM62 alignment score matrix come from? *Nat. Biotechnol.* 22, 1035–1036. doi:10.1038/nbt0804-1035

Einav, T., and Phillips, R. (2019). How the avidity of polymerase binding to the −35/−10 promoter sites affects gene expression. *Proc. Natl. Acad. Sci.* 116, 13340–13345. doi:10.1073/pnas.1905615116

Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., et al. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol.* 35, 48–55. doi:10.1038/nbt.3718

Gebauer, F., Preiss, T., and Hentze, M. W. (2012). From cis-regulatory elements to complex RNPs and back. *Cold Spring Harb. Perspect. Biol.* 4, a012245. doi:10.1101/cshperspect.a012245

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi:10.1038/nmeth.1318

Godara, A., and Kao, K. C. (2020). Adaptive laboratory evolution for growth coupled microbial production. *World J. Microbiol. Biotechnol.* 36, 175. doi:10.1007/s11274-020-02946-8

Harrington, L. B., Jha, R. K., Kern, T. L., Schmidt, E. N., Canales, G. M., Finney, K. B., et al. (2017). Rapid thermostabilization of *Bacillus thuringiensis* serovar konkukian 97–27 dehydroshikimate dehydratase through a structure-based enzyme design and whole cell activity assay. *ACS Synth. Biol.* 6 (1), 120–129. doi:10.1021/acssynbio.6b00159

Hartman, A. H., Liu, H., and Melville, S. B. (2011). Construction and characterization of a lactose-inducible promoter system for controlled gene expression in *Clostridium perfringens*. *Appl. Environ. Microbiol.* 77, 471–478. doi:10.1128/aem.01536-10

Heckman, K. L., and Pease, L. R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* 2, 924–932. doi:10.1038/nprot.2007.132

Henikoff, S., and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A.* 89, 10915–10919. doi:10.1073/pnas.89.22.10915

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77, 51–59. doi:10.1016/0378-1119(89)90358-2

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. *Gene* 77, 61–68. doi:10.1016/0378-1119(89)90359-4

Jajesniak, P., and Wong, T. S. (2015). QuickStep-cloning: A sequence-independent, ligation-free method for rapid construction of recombinant plasmids. *J. Biol. Eng.* 9, 15. doi:10.1186/s13036-015-0010-3

Jha, R. K., Bingen, J. M., Johnson, C. W., Kern, T. L., Khanna, P., Trettel, D. S., et al. (2018). A protocatechuate biosensor for Pseudomonas putida KT2440 via promoter and protein evolution. *Metab. Eng. Commun.* 6, 33–38. doi:10.1016/j.meteno.2018.03.001

Jha, R. K., and Strauss, C. E. M. (2020). Smart microbial cells couple catalysis and sensing to provide high-throughput selection of an organophosphate hydrolase. *ACS Synth. Biol.* 9 (6), 1234–1239. doi:10.1021/acssynbio.0c00025

Jha, R. K., Chakraborti, S., Kern, T. L., Fox, D. T., and Strauss, C. E. M. (2015). Rosetta comparative modeling for library design: Engineering alternative inducer specificity in a transcription factor. *Proteins Struct. Funct. Bioinforma.* 83, 1327–1340. doi:10.1002/prot.24828

Jha, R. K., Kern, T. L., Fox, D. T., and Strauss, C. E. M. (2014). Engineering an Acinetobacter regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry. *Nucleic Acid. Res.* 42, 8150–8160. doi:10.1093/nar/gku444

Jha, R. K., Kern, T. L., Kim, Y., Tesar, C., Jedrzejczak, R., Joachimiak, A., et al. (2016). A microbial sensor for organophosphate hydrolysis exploiting an engineered specificity

switch in a transcription factor. *Nucleic Acids Res.* 44, 8490–8500. doi:10.1093/nar/gkw687

Jha, R. K., Narayanan, N., Pandey, N., Bingen, J. M., Kern, T. L., Johnson, C. W., et al. (2019). Sensor-enabled alleviation of product inhibition in chorismate pyruvate-lyase. *ACS Synth. Biol.* 8, 775–786. doi:10.1021/acssynbio.8b00465

Jin, L.-Q., Jin, W. R., Ma, Z. C., Shen, Q., Cai, X., Liu, Z. Q., et al. (2019). Promoter engineering strategies for the overproduction of valuable metabolites in microbes. *Appl. Microbiol. Biotechnol.* 103, 8725–8736. doi:10.1007/s00253-019-10172-y

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2

Kosuri, S., and Church, G. M. (2014). Large-scale de novo DNA synthesis: Technologies and applications. *Nat. Methods* 11, 499–507. doi:10.1038/nmeth.2918

Kretz, K. A., Richardson, T. H., Gray, K. A., Robertson, D. E., Tan, X., Short, J. M., et al. (2004). "Gene site saturation mutagenesis: A comprehensive mutagenesis approach," in *Methods in enzymology* (Academic Press), 388, 3–11.

LaFleur, T. L., Hossain, A., and Salis, H. M. (2022). Automated model-predictive design of synthetic promoters to control transcriptional profiles in bacteria. *Nat. Commun.* 13, 5159. doi:10.1038/s41467-022-32829-5

Li, K., Shrivastava, S., Brownley, A., Katzel, D., Bera, J., Nguyen, A. T., et al. (2012). Automated degenerate PCR primer design for high-throughput sequencing improves efficiency of viral sequencing. *Virol. J.* 9, 261. doi:10.1186/1743-422x-9-261

Lynch, M., and Hill, W. G. (1986). Phenotypic evolution by neutral mutation. *Evolution* 40, 915–935. doi:10.2307/2408753

Lyskov, S., Chou, F. C., Conchúir, S. Ó., Der, B. S., Drew, K., Kuroda, D., et al. (2013). Serverification of molecular modeling applications: The Rosetta online server that includes everyone (ROSIE). *PLOS ONE* 8, e63906. doi:10.1371/journal.pone.0063906

Lyskov, S., and Gray, J. J. (2008). The RosettaDock server for local protein–protein docking. *Nucleic Acids Res.* 36, W233–W238. doi:10.1093/nar/gkn216

Maddocks, S. E., and Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiol. Read. Engl.* 154, 3609–3623. doi:10.1099/mic.0.2008/022772-0

Maity, T. S., Jha, R. K., Strauss, C. E. M., and Dunbar, J. (2012). Exploring the sequence–function relationship in transcriptional regulation by the lac O1 operator. *FEBS J.* 279, 2534–2543. doi:10.1111/j.1742-4658.2012.08635.x

The UniProt Consortium, Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., et al. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. doi:10.1093/nar/gkaa1100

Mayo, A. E., Setty, Y., Shavit, S., Zaslaver, A., and Alon, U. (2006). Plasticity of the cis-regulatory input function of a gene. *PLOS Biol.* 4, e45. doi:10.1371/journal.pbio.0040045

Miyazaki, K. (2011). "Chapter seventeen - MEGAWHOP cloning: A method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids," in *Methods in enzymology*. Editor C. Voigt (Academic Press), 498, 399–406.

Miyazaki, K. (2011). "MEGAWHOP cloning," in *Methods in enzymology* (Elsevier), 498, 399–406.

Miyazaki, K., and Takenouchi, M. (2002). Creating random mutagenesis libraries using megaprimer PCR of whole plasmid. *Biotechniques* 33, 1033–1038. doi:10.2144/02335st03

Mordaka, P. M., and Heap, J. T. (2018). Stringency of synthetic promoter sequences in *Clostridium* revealed and circumvented by tuning promoter library mutation rates. *ACS Synth. Biol.* 7, 672–681. doi:10.1021/acssynbio.7b00398

Oesterle, S., Gerngross, D., Schmitt, S., Roberts, T. M., and Panke, S. (2017). Efficient engineering of chromosomal ribosome binding site libraries in mismatch repair proficient *Escherichia coli*. *Sci. Rep. Nat. Publ. Group* 7, 12327. doi:10.1038/s41598-017-12395-3

Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., et al. (2019). Lessons from two design–build–test–learn cycles of dodecanol production in *Escherichia coli* aided by machine learning. *ACS Synth. Biol.* 8, 1337–1351. doi:10.1021/acssynbio.9b00020

Pandey, N., Davison, S. A., Krishnamurthy, M., Trettel, D. S., Lo, C. C., Starkenburg, S., et al. (2022). Precise genomic riboregulator control of metabolic flux in microbial systems. *ACS Synth. Biol.* 11, 3216–3227. doi:10.1021/acssynbio.1c00638

Pardo, I., Jha, R. K., Bermel, R. E., Bratti, F., Gaddis, M., McIntyre, E., et al. (2020). Gene amplification, laboratory evolution, and biosensor screening reveal MucK as a terephthalic acid transporter in *Acinetobacter baylyi* ADP1. *Metab. Eng.* 62, 260–274. doi:10.1016/j.ymben.2020.09.009

Peccoud, J. (2016). Synthetic biology: Fostering the cyber-biological revolution. *Synth. Biol.* 1, ysw001. doi:10.1093/synbio/ysw001

Rhodius, V. A., Mutalik, V. K., and Gross, C. A. (2012). Predicting the strength of UP-elements and full-length *E. coli* σE promoters. *Nucleic Acids Res.* 40, 2907–2924. doi:10.1093/nar/gkr1190

Rocklin, G. J., Chidyausiku, T. M., Goreshnik, I., Ford, A., Houliston, S., Lemak, A., et al. (2017). Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* 357, 168–175. doi:10.1126/science.aan0693

Rohl, C. A., Strauss, C. E. M., Misura, K. M. S., and Baker, D. (2004). "Protein structure prediction using Rosetta," in *Methods in enzymology* (Academic Press), 383 66–93.

Saecker, R. M., Record, M. T., and deHaseth, P. L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.* 412, 754–771. doi:10.1016/j.jmb.2011.01.018

Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27, 946–950. doi:10.1038/nbt.1568

Sayous, V., Lubrano, P., Li, Y., and Acevedo-Rocha, C. G. (2020). Unbiased libraries in protein directed evolution. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 1868, 140321. doi:10.1016/j.bbapap.2019.140321

Segall-Shapiro, T. H., Sontag, E. D., and Voigt, C. A. (2018). Engineered promoters enable constant gene expression at any copy number in bacteria. *Nat. Biotechnol.* 36, 352–358. doi:10.1038/nbt.4111

She, W., Ni, J., Shui, K., Wang, F., He, R., Xue, J., et al. (2018). Rapid and error-free site-directed mutagenesis by a PCR-free *in vitro* CRISPR/Cas9-Mediated mutagenic system. *ACS Synth. Biol.* 7, 2236–2244. doi:10.1021/acssynbio.8b00245

Shin, S.-M., Jha, R. K., and Dale, T. (2022). Tackling the catch-22 situation of optimizing a sensor and a transporter system in a whole-cell microbial biosensor design for an anthropogenic small molecule. *ACS Synth. Biol.* 11, 3996–4008. doi:10.1021/acssynbio.2c00364

Siloto, R. M. P., and Weselake, R. J. (2012). Site saturation mutagenesis: Methods and applications in protein engineering. *Biocatal. Agric. Biotechnol.* 1, 181–189. doi:10.1016/j.bcab.2012.03.010

Tarasava, K., Oh, E. J., Eckert, C. A., and Gill, R. T. (2018). CRISPR-enabled tools for engineering microbial genomes and phenotypes. *Biotechnol. J.* 13, 1700586. doi:10.1002/biot.201700586

Thomas, S., Maynard, N. D., and Gill, J. (2015). DNA library construction using Gibson Assembly®. *Nat. Methods* 12, i–ii. doi:10.1038/nmeth.f.384

Thomason, L. C., Sawitzke, J. A., Li, X., Costantino, N., and Court, D. L. (2014). Recombineering: Genetic engineering in bacteria using homologous recombination. *Curr. Protoc. Mol. Biol.* 106, 1.16.1–1.16.39. doi:10.1002/0471142727.mb0116s106

Waller, M. C., Bober, J. R., Nair, N. U., and Beisel, C. L. (2017). Toward a genetic tool development pipeline for host-associated bacteria. *Curr. Opin. Microbiol.* 38, 156–164. doi:10.1016/j.mib.2017.05.006

White, M. A. (2015). Understanding how cis-regulatory function is encoded in DNA sequence using massively parallel reporter assays and designed sequences. *Genomics* 106, 165–170. doi:10.1016/j.ygeno.2015.06.003

Wong, T. S., Tee, K. L., Hauer, B., and Schwaneberg, U. (2004). Sequence saturation mutagenesis (SeSaM): A novel method for directed evolution. *Nucleic Acids Res.* 32, e26–e26. doi:10.1093/nar/gnh028

Xie, M., and Fussenegger, M. (2018). Designing cell function: Assembly of synthetic gene circuits for cell biology applications. *Nat. Rev. Mol. Cell. Biol.* 19, 507–525. doi:10.1038/s41580-018-0024-z

Xiong, W., Liu, B., Shen, Y., Jing, K., and Savage, T. R. (2021). Protein engineering design from directed evolution to de novo synthesis. *Biochem. Eng. J.* 174, 108096. doi:10.1016/j.bej.2021.108096

Yang, J., Ruff, A. J., Arlt, M., and Schwaneberg, U. (2017). Casting epPCR (cepPCR): A simple random mutagenesis method to generate high quality mutant libraries. *Biotechnol. Bioeng.* 114, 1921–1927. doi:10.1002/bit.26327

Yin, J., Zhu, Y., Liang, Y., Luo, Y., Lou, J., Hu, X., et al. (2022). Development of whole-cell biosensors for screening of peptidoglycan-targeting antibiotics in a gram-negative bacterium. *Appl. Environ. Microbiol.* 88, 00846-22. doi:10.1128/aem.00846-22

Zhang, B., Zhou, N., Liu, Y. M., Liu, C., Lou, C. B., Jiang, C. Y., et al. (2015). Ribosome binding site libraries and pathway modules for shikimic acid synthesis with *Corynebacterium glutamicum*. *Microb. Cell. Factories* 14, 71. doi:10.1186/s12934-015-0254-0

Zhang, X., Xu, G., Shi, J., and Xu, Z. (2018). Integration of ARTP mutagenesis with biosensor-mediated high-throughput screening to improve l-serine yield in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 102, 5939–5951. doi:10.1007/s00253-018-9025-2

Zheng, Y., Hong, K., Wang, B., Liu, D., Chen, T., and Wang, Z. (2021). Genetic diversity for accelerating microbial adaptive laboratory evolution. *ACS Synth. Biol.* 10, 1574–1586. doi:10.1021/acssynbio.0c00589

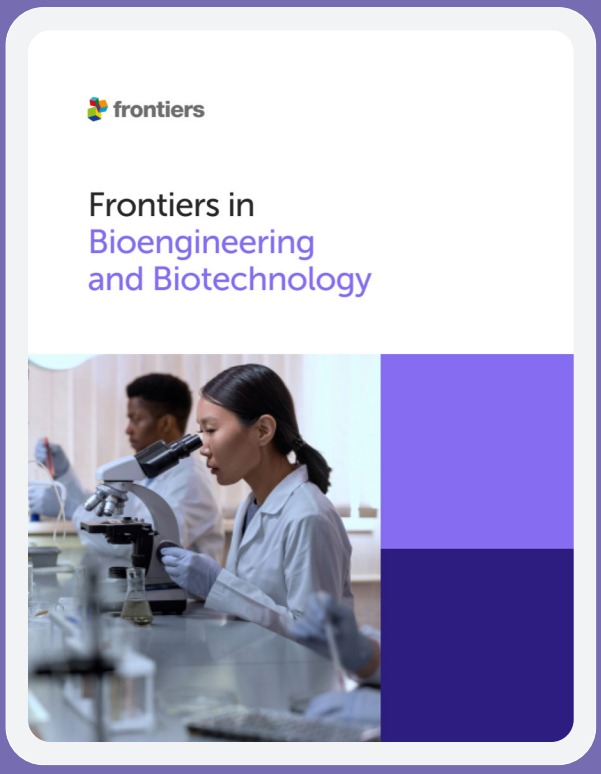
Frontiers in
Bioengineering and Biotechnology

Accelerates the development of therapies, devices, and technologies to improve our lives

A multidisciplinary journal that accelerates the development of biological therapies, devices, processes and technologies to improve our lives by bridging the gap between discoveries and their application.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact