



Genomic and transcriptional changes in response to pinene tolerance and overproduction in evolved *Escherichia coli*

Fu-Xing Niu, Yuan-Bin Huang, Liang-Nian Ji**, Jian-Zhong Liu*

Institute of Synthetic Biology, MOE Key Laboratory of Bioinorganic and Synthetic Chemistry, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China



ARTICLE INFO

Keywords:
 Comparative omics analysis
 CRISPRa
 CRISPRi
 Pinene
Escherichia coli
 Biosynthesis

ABSTRACT

α -Pinene is an important monoterpene, which is widely used as a flavoring agent and in fragrances, pharmaceuticals and biofuels. Although an evolved strain *Escherichia coli* YZFP, which had higher tolerance to pinene and titer, has been successfully used to produce high levels of pinene, the pinene titer is much lower than that of hemiterpene (isoprene) and sesquiterpenes (farnesene) to date. Moreover, the overall cellular physiological and metabolic changes caused by higher tolerance to pinene and overproduction of pinene remains unclear. To reveal the mechanism of *Escherichia coli* YZFP with the higher tolerance to pinene and titer, a comparative genomics and transcriptional level analyses combining with CRISPR activation (CRISPRa) and interference (CRISPRi) were carried out. The results show that the tolerance to pinene and the overproduction of pinene in *E. coli* may be associated with: 1) the mutations of the DXS pathway genes, the *rpoA* and some membrane protein genes, and their upregulations of transcription levels; and 2) the mutations of some genes and their down-regulation of transcriptional levels. These comparative omics analyses provided some genetic modification strategies to further improve pinene production. Overexpression of the mutated *cbpA*, *tabA*, *pitA*, *rpoA*, *sufBCDS*, *mutS*, *ispH*, *oppF*, *dusB*, *dnaK*, *dxs*, *dxr* and *flgFGH* genes further improved pinene production. This study also demonstrated that combining comparative omics analysis with CRISPRa and CRISPRi is an efficient technology to quickly find a new metabolic engineering strategy.

1. Introduction

α -Pinene is a natural and active monoterpene, which is widely used as flavoring agents [1], chemicals [2] and pharmaceuticals [3], especially as a good candidate for jet fuel replacement. As a result, the biotechnological production of pinene has received increasing attention. Yang et al. first engineered an *E. coli* to produce 5.4 mg/L pinene by the coexpression of a heterologous mevalonate (MEV) pathway and α -pinene synthase (Pt 30) from *Pinus taeda* [4]. Then, the combinatorial expression of *Abies grandis* geranyl diphosphate synthase (GPPS)-pinene synthase (PS) fusion proteins further enhanced pinene production to 32 mg/L in *E. coli* [5]. The directed evolution of α -pinene synthase (Pt1) from *P. taeda* resulted in α -pinene production to 140 mg/L in *E. coli* [6]. Dunlop et al. reported that *E. coli* cannot grow in the presence of 0.5% [7]. To avoid high toxicity of pinene on the growth of *E. coli* and biosynthesis of pinene, we first improved pinene tolerance to 2.0% and pinene production to 9.9 mg/L from 5.6 mg/L through

adaptive laboratory evolution after atmospheric and room temperature plasma (ARTP) mutagenesis and overexpression of the efflux pump to obtain the pinene tolerant strain *Escherichia coli* YZFP [8]. Then, we applied a combinatorial strategy of enzymatic directed evolution and modular co-culture engineering to further increase pinene production to 166.5 mg/L [8]. In addition, *Corynebacterium glutamicum* has also been successfully used as host for pinene biosynthesis. Coexpression of native 1-deoxy-d-xylulose-5-phosphate synthase (Dxs), isopentenyl diphosphate isomerase (Idi), *P. taeda* PS and *A. grandis* GPPS in *C. glutamicum* yielded pinene of 27 μ g/g cell dry weight [9]. Although some progress has been made in pinene biosynthesis, the production level of pinene is much lower than that of hemiterpene (isoprene) (gram level) [8,10]. Thus, further work is required to increase the production of pinene.

In this study, we combined comparative genomics and transcriptional level analysis with CRISPR activation (CRISPRa) and inference (CRISPRi) to investigate the mechanism of *E. coli* YZFP with higher

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

* Corresponding author.

** Corresponding author.

E-mail addresses: niuwx3@mail.sysu.edu.cn (F.-X. Niu), huangyb35@mail2.sysu.edu.cn (Y.-B. Huang), cesjln@mail.sysu.edu.cn (L.-N. Ji), lssljz@mail.sysu.edu.cn (J.-Z. Liu).

<https://doi.org/10.1016/j.synbio.2019.05.001>

Received 19 March 2019; Received in revised form 19 May 2019; Accepted 19 May 2019

2405-805X/© 2019 Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1
Strains and plasmids used in this study.

Name	Description	Sources
Strain		
<i>E. coli</i> DH5 α	F- <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG</i> Φ 80 <i>dlacZ</i> Δ M15 Δ (<i>lacZYA-argF</i>) U169	Invitrogen
<i>E. coli</i> BW25113(P _{T5-dxs})	<i>lacI^q rrrB_{T14} ΔlacZ_{WJ16} hsdR514 ΔsarA AD_{AH33} ΔrhaBAD_{LD78}</i> , P _{dxc} ::P _{T5}	[11]
<i>E. coli</i> YZFP	The adaptive laboratory evolution strain from <i>E. coli</i> BW25113 (P _{T5-dxs}), P _{acrAB} ::P ₃₇ , SS9::P ₃₇ -ttgB, tolerance to 2.0% pinene and 35 μ M fosmidomycin	[8]
Plasmid		
<i>E. coli</i> PINE	Pinene producer, chemically induced chromosomal evolution (CICHe) strain from <i>E. coli</i> <i>E. coli</i> YZFP after integration of the TIGR-mediated gene cluster of the <i>Abies grandis</i> GPPS ^{Mut} . <i>Pinus taeda</i> Pt1 ^{MUT} gene cluster	[8]
<i>E. coli</i> MEVI	CICHe strain from <i>E. coli</i> YZFP after integration of the mevalonate pathway	[8]
pMEVIGPS	pBbA5C-MTSAe-T1f-MBI(f)-T1002i-trGPPS _{A.grandis} -PS _{A.grandi} coding for MEV pathway enzymes to produce pinene from glucose in <i>E. coli</i> , p1FA <i>ori</i> , P _{lacUVF} promoter, Cm ^r	[8]
pU57-RMS	pUC57 harboring P _{tet} promoter and scrRNA sequences	This study
pCRISPathBrick*	Expressing <i>dcas9</i> (K848A/K1003A/R1060A), p15A <i>ori</i> , Cm ^r	[12,13]
pBbB2K-GFP	Addgene plasmid #35345, P _{tet} promoter, Kan ^r , pBBR1 <i>ori</i>	[14]
pBbB2K-GFP*	pBbB2K-GFP harboring <i>XhoI</i> endonuclease site after tet promoter	This study
pBbB2K-dCas9*	pBbB2K containing the <i>dcas9</i> (K848A/K1003A/R1060A)	This study
pUC57-MCPSoxS	pUC57 harboring the fusion gene of the codon-optimized MCP and SOX with a (GGGS) linker, P _{tet} Promoter	This study
pBbB2K-dCas9*-MCPSoxS	pBbB2K-dCas9* containing the MCPSoxS sequences	This study
pTargetB	<i>E. coli</i> sgRNA expression vector, BglBrick vector, pMB1, Spe ^r , scrRNA	[8]
pTargetA	<i>E. coli</i> scrRNA expression vector, BglBrick vector, P _{tet} promoter, Spe ^r , pMB1 <i>ori</i>	This study
pZEABP	Constitute expression vector, pBR322 <i>ori</i> , P37 promoter, Amp ^r	[15]

tolerance to pinene and overproduction. Based on these results, over-expression of the mutants further improved pinene production.

2. Materials and methods

2.1. Strains, plasmids and primers

The bacterial strains and plasmids used in this study are listed in Table 1. The primers used in this study are listed in Supplementary Table 1.

2.2. Whole genome resequencing and data analyses

Genomic DNA from *E. coli* YZFP strain was extracted using TIANamp Bacterial DNA Kit (Tiangen Biotech Co., Beijing, China) according to the manufacturer's instruction. Genomic library construction and whole-genome resequencing were performed on the Illumina HiSeq 2500 sequencing platform by GENWIZ (Shuzhou, China). The paired-end reads from *E. coli* YZFP were aligned to the reference genome of *E. coli* MG1655 using the BWA software (version 0.7.12). Potential mutations including point mutation and insertion/deletions were identified using the Samtools software (version 1.1) and GATK module (Unified Genotyper).

2.3. Transcriptional level analysis

The total RNA from *E. coli* cells grown for 30 h in shake flasks was isolated using an RNA extraction kit (Dongsheng Biotech, Guangzhou, China), following the manufacturer's instructions. RNase-free gDNaseI (Dongsheng Biotech, Guangzhou, China) was treated during the isolation procedure to eliminate possible DNA contamination. Absence of DNA was verified by control PCR reactions using the RNA as a template. RNA concentration was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific). Only the RNA samples with 260/280 ratio (an indication of protein contamination) greater than 1.8 and 260/230 ratio (an indication of reagent contamination) greater than 2.0 were used for the analysis. The first-strand cDNA was synthesized using an All-in-One™ First-Strand cDNA Synthesis kit (GeneCopia, Guangzhou, China). Quantitative real-time PCR (qRT-PCR) was performed with the All-in-One™ qPCR Mix kit (GeneCopia) by an iCycler iQ5 Real Time PCR system (Bio-Rad Laboratories, California, USA). The template was 100 ng of cDNA. The PCR conditions were as

follows: 95 °C for 10 min, followed by 45 cycles of denaturation at 95 °C for 10 s, annealing at 60 °C for 20 s, and extension at 72 °C for 15 s. The primers for qRT-PCR are presented in Supplementary Table 1. All qRT-PCR quantifications were performed simultaneously with PCR amplified standard and no-template (water) controls. The expression levels were analyzed by the $2^{-\Delta\Delta Ct}$ method described by Livak and Schmittgen [16] and normalized by *cysG* gene expression. Finally, the ratio of the transcriptional level of the evolved *E. coli* YZFP strain and original *E. coli* BW25113 (P_{T5-dxs}) strain of > 1.2 or < 0.8 ($P < 0.05$, Student's t-test) was considered significant differential expression. Three biological replicates for each sample were used for qRT-PCR analysis and three technical replicates were analyzed for each biological replicate.

2.4. CRISPRa and CRISPRi of the mutant genes

Dong et al. developed a CRISPR-Cas system based on the actor SoxS with simultaneous activation and repression [17]. We applied this CRISPR-Cas-SoxS system [17] for activating and repressing target genes in *E. coli* BW25113(P_{T5-dxs}). To enhance specificity and reduce off-target effects, *dcas9** (K848A/K1003A/R1060A) was amplified from pCRIS-PathBrick* with primers *dcas9**-F/*dcas9**-R and then cloned into the *ApaI/XhoI* sites of pBbB2k-GFP* to generate pBbB2k-dCas9*. P_{tet}-MCPSoxS fragment was synthesized by GENWIZ (Shuzhou, China) and ligate into pUC57 to obtain pUC57-MCPSOX. This fragment was cut from pUC57-MCPSOX with *BamHI/XhoI*, and then cloned into *BglIII/XhoI*-digested pBbB2K-dCas9* to obtain pBbB2K-dCas9*-MCPSoxS (Supplementary Fig. 1A).

Scaffold RNA (scrRNA) expressing vector pTargetA was derived from pTargetB [8]. P_{tet}-scrRNA fragment was synthesized by GENWIZ (Shuzhou, China) and ligate into pUC57 to obtain pUC57-RMS. This fragment was cut from pUC57-RMS with *NdeI/EcoRI*, and then cloned into *NdeI/EcoRI*-digested pTargetB to obtain pTargetA.

The pTargetA-X series used in target single-gene activation or repression with a targeting N20 sequence of gene loci of interest, was obtained by inverse PCR using primers the target N20F and N20-genes from pTargetA, then cut with *KpnI* and self-ligation. For activation, the N20 sequence was designed to target at the non-template strand at the upstream of the promoter. For repression, The N20 sequence was designed to target at the 5' end (about 100 bp downstream of ATG) of the gene on the non-template DNA strand. Because the scrRNA fragment was flanked with *BamHI* and *BglIII*, the scrRNA expressing vector pTargetA-

Table 2
The SNV and InDel mutants in *E. coli* YZFP^a.

Type	Amounts of mutant	Amounts of gene
Synonymous SNV	93	84
Nonsynonymous SNV	184	175
Stop gaining SNV	11	11
Intergenic SNV	2	
Upstream SNV	14	
Downstream SNV	15	
Upstream and Downstream SNV	29	
Total	349	270
InDel	25	

^a Mapped reads ratio of 99.79%, covered bases of 4612119 bp, unique read ratio of 97.77%, genome coverage of 99.36%, mean depth of 399X. The SNVs and InDels were identified with a frequency of 1.0.

XY for multiple genes was constructed from pTargetA-X and pTargetA-Y by using the standard BglBrick assemble approach [14].

To investigate the effect of the activation and repression on growth, pBbB2K-dCas9*-MCPSoxS and pTargetA-X were co-transferred into *E. coli* BW25113(P_{T5-dxs}). A single colony was grown in 5 mL LB medium in a falcon tube at 37 °C overnight. The overnight cultures were inoculated into 50 mL of LB medium supplement with 0.5% pinene and 200 nM anhydrous tetracycline with a starting OD₆₀₀ of 0.09. The cultures were incubated at 37 °C (200 rpm) for 12 h. The strain harboring the empty vector pTargetA, *E. coli* BW25113(P_{T5-dxs}) (pBbB2K-dCas9*-MCPSoxS, pTargetA), was set as the control.

To investigate the effect of the activation and repression on pinene production, pBbB2K-dCas9*-MCPSoxS, pTargetA-X and pMEVIGPS were co-transferred into *E. coli* BW25113(P_{T5-dxs}) and then analyzed the production of pinene in shake flasks. The strain harboring the empty vector pTargetA, *E. coli* BW25113(P_{T5-dxs}) (pBbB2K-dCas9*-MCPSoxS, pTargetA, pMEVIGPS) was set as the control.

2.5. Pinene production in shake flasks

A single colony was inoculated into 5 mL of LB medium in a falcon tube, which was cultured overnight at 37 °C and 200 rpm. The overnight seed culture was then inoculated into 50 mL of SBMSN medium with a starting OD₆₀₀ of 0.1. SBMSN medium (pH 7.0) containing the following (g/L): sucrose 20, peptone 12, yeast extract 24, KH₂PO₄ 1.7, K₂HPO₄ 211.42, MgCl₂·6H₂O 1, ammonium oxalate 1.42, and Tween-80 2 mL. The main cultures were then incubated at 37 °C and 200 rpm until an OD₆₀₀ of 0.8 was reached. Then, the cultures were induced with 1 mM IPTG and 200 nM anhydrous tetracycline, and overlaid with 20%

dodecane to trap pinene. After induction, the cultures were incubated at 30 °C and 130 rpm for 72 h.

2.6. Whole-cell biocatalysis for pinene production

Genes were amplified from *E. coli* YZFP genomes using the responding primers (Supplementary Table 1) and cloned into pZEABP to obtain pZEA-gene series plasmids.

E. coli PINE harboring the target plasmid pZEA-gene and *E. coli* MEVI harboring the pZEABP were used for whole-cell biocatalysis. They were separately inoculated into 5 mL of SBMSN medium in a falcon tube, which was cultured overnight at 37 °C. The overnight cultures were then inoculated into 50 mL SBMSN medium with a starting OD₆₀₀ of 0.1. The cultures were incubated at 20 °C for 16 h. The *E. coli* PINE culture was mixed with *E. coli* MEVI culture at the inoculation ration of 2:1. The mixed cells were harvested by centrifugation (6000 × g at 4 °C) and washed twice with cooled phosphate buffer (0.1 M, pH 7.0). Finally, the above cells were resuspended in 10 mL phosphate buffer (0.1 M, pH 7.0) containing 20 g/L of sucrose, 10 mM MgCl₂ and 5 mM MnCl₂ to form the cell suspension (OD₆₀₀ = 30). The reaction mixture was overlaid with 20% dodecane. The catalysis was performed for 28 h at 30 °C and 130 rpm.

For inverse engineering using CRISPRi, *E. coli* PINE (pBbB2K-dCas9*-MCPSoxS, pTargetA-X) and *E. coli* MEVI (pBbB2K-GFP, pTargetA) were used for whole-cell biocatalysis as the above-mentioned process.

2.7. GC analysis

Pinene concentration in the dodecane layer was assayed as previously reported [8]. Take the dodecane layer and dilute it in ethylacetate spiked with the internal standard limonene. The samples were analyzed by GC-FID by using a standard curve of α -pinene (Sigma Aldrich). The GC-FID (Techcomp GC7900, Techcomp Ltd, China) was used with a TM-5 column (30 m × 0.32 mm × 0.50 μ m). The inlet temperature was set to 300 °C, with a flow at 1 mL/min, oven at 50 °C for 30 s, ramp at 4 °C/min to 70 °C, and ramp at 25 °C/min to 240 °C.

2.8. Statistical analysis

All experiments were conducted in triplicate, and data were averaged and presented as the means \pm standard deviation. Oneway analysis of variance followed by Tukey's test was used to determine significant differences using the OriginPro (version 9.0) package. Statistical significance was defined as $p < 0.05$.

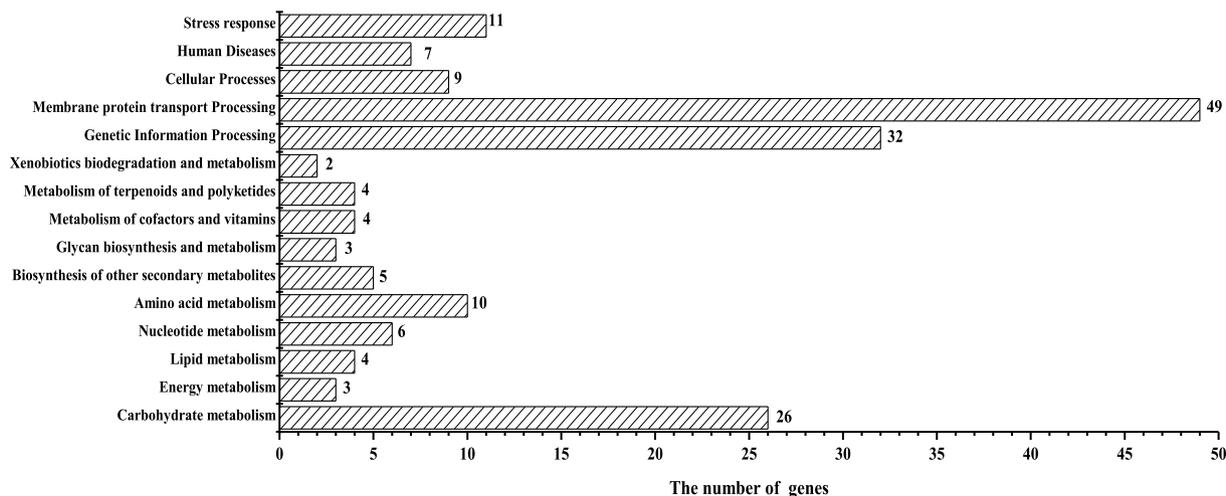


Fig. 1. Pathway gene ontology enrichment analysis of the nonsynonymous mutant genes.

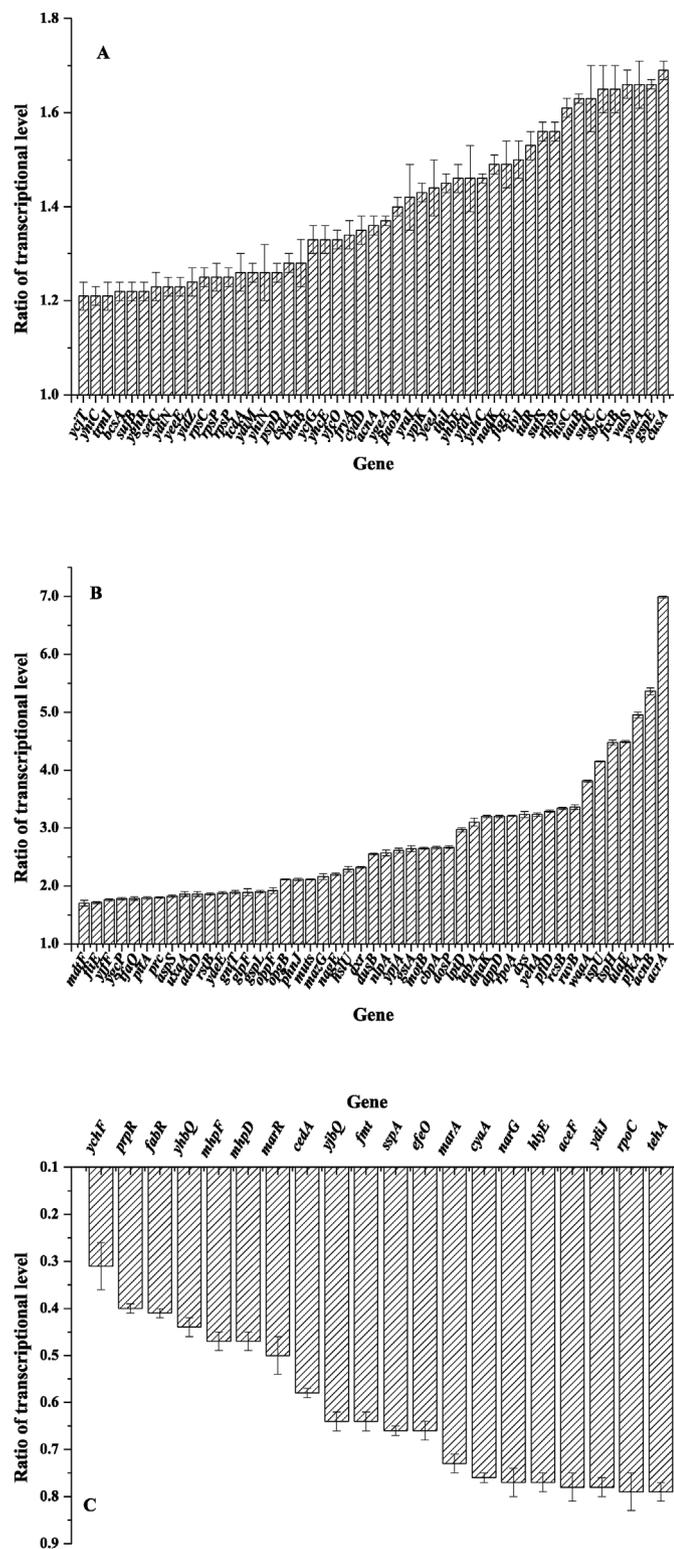


Fig. 2. The relative transcriptional levels of the nonsynonymously mutated genes in *E. coli* YZFP compared with those in *E. coli* BW25113(P_{T5-dxs}). (A) Ratio was less than 1.7; (B) ratio was greater than 1.7; (C) ratio was less than 0.8. The data represent the means of three replicates, and error bars represent standard deviations.

3. Results and discussion

3.1. Whole genome resequencing

In our previous study [8], we obtain an evolved strain *E. coli* YZFP, which showed a higher tolerance to pinene and titer. To investigate the relationship between the genotype and the phenotype of *E. coli* YZFP, the whole genome of *E. coli* YZFP was re-sequenced using a 2×150 paired-end (PE) configuration. The genome sequencing generated 13548860 clean reads and achieved a mean depth of 399× and the mapped reads ratio of 99.79% (Table 2). After alignment to the genome sequence of *E. coli* MG1655, a genome coverage of 99.36% with covered bases of 4612119 bp was obtained. A total of 349 single nucleotide variants (SNV) and 25 insertion/deletion (InDel) were identified. Of SNVs, 184 nonsynonymous mutation distributed among 175 mutated genes (Table 2 and Supplementary Table 2). These nonsynonymous mutant genes were mainly involved in the membrane protein transport processing (49 genes), the genetic information processing (32 genes), the carbohydrate metabolism (26 genes), the stress response (11 genes), the amino acid metabolism (10 genes), the cellular processes (9 genes), the human diseases (7 genes), the nucleotide metabolism (6 genes), the biosynthesis of other secondary metabolism (5 genes), the lipid metabolism (4 genes), the metabolism of cofactors and vitamins (4 genes), the metabolism of terpenoids and polyketides (4 genes), the glycan biosynthesis and metabolism (3 genes), the energy metabolism (3 genes), and the xenobiotics biodegradation and metabolism (2 genes) (Fig. 1).

3.2. Comparative transcriptional level analysis of the mutant genes

Although comparative transcriptome analysis can provide more information, it is difficult to confirm that transcriptome changes are caused by gene mutations. Thus, we only analyzed the changes in transcriptional levels of the mutant genes. To investigate whether gene mutations caused changes in transcriptional levels in *E. coli* YZFP, the transcriptional levels of the nonsynonymous mutant genes (174 genes) and genes observed SNV in the upstream (8 genes) were determined using qRT-PCR and were compared with those in *E. coli* BW25113(P_{T5-dxs}). A total of 116 genes showed significantly different expression patterns. Of them, 96 genes were up-regulated and 20 genes were down-regulated in *E. coli* YZFP compared to *E. coli* BW25113(P_{T5-dxs}) (Fig. 2).

3.3. Functional confirmation of the mutant genes

Dong et al. developed a CRISPR-Cas system based on the actor SoxS with simultaneous activation and repression [17]. Using this CRISPR-Cas-SoxS system enable programmable control over multiple genes with simultaneous activation and repression. Thus, we first constructed a dcas9* and SoxS expressing plasmid pBbB2K-dCas9*-MCPSoxS and scRNA expressing vector pTargetA (Supplementary Fig. 1A). Then, we characterized the functions of the CRISPR-Cas-SoxS system. We constructed the scRNA plasmid for CRISPRa and CRISPRi to mCherry, respectively. When N20 sequence was designed to target the P37 promoter or the CDS of mCherry, mCherry was repressed (Supplementary Fig. 1B). When N20 sequence was designed to target the upstream of the P37 promoter, mCherry was activated (Supplementary Fig. 1C). These results demonstrate that the CRISPR-Cas-SoxS system can function with activation and repression.

Thus, we applied this CRISPR-Cas-SoxS system [17] for activating and repressing target genes in *E. coli* BW25113(P_{T5-dxs}) to investigate their effects on growth and pinene production. Of 96 up-regulated genes, a total of 57 up-regulated genes with higher transcription level compared to *E. coli* BW25113(P_{T5-dxs}) was selected to be activated with CRISPRa. pBbB2K-dCas9*-MCPSoxS and pTargetA-X were co-transferred into *E. coli* BW25113(P_{T5-dxs}). The growth and pinene production

Table 3Effect of CRISPR activation of the up-regulated genes on growth and production of pinene in *E. coli* BW25113(P_{T5-dxs}).

Gene	Protein	Ratio of growth ^a	Ratio of pinene concentration ^b
Carbohydrate metabolism			
<i>ulaE</i>	L-xylulose 5-phosphate 3-epimerase	1.24 ± 0.01	1.06 ± 0.02
Metabolism of terpenoids and polyketides			
<i>ispH</i>	4-hydroxy-3-methylbut-2-enyl diphosphate reductase, 4Fe-4S protein	1.07 ± 0.01	1.19 ± 0.01
<i>dxr</i>	1-deoxy-D-xylulose 5-phosphate reductoisomerase	–	1.34 ± 0.02
<i>dxs</i>	1-deoxyxylulose-5-phosphate synthase, thiamine triphosphate-binding, FAD-requiring	–	1.39 ± 0.02
<i>ispU</i>	1-deoxy-D-xylulose 5-phosphate reductoisomerase/undecaprenyl pyrophosphate synthase	–	1.28 ± 0.01
Genetic Information Processing			
<i>rpoA</i>	RNA polymerase, alpha subunit	1.50 ± 0.01	1.09 ± 0.01
<i>dosP</i>	oxygen sensor, c-di-GMP phosphodiesterase,	1.56 ± 0.02	1.07 ± 0.01
<i>mutS</i>	methyl-directed mismatch repair protein	1.10 ± 0.01	1.29 ± 0.01
<i>cbpA</i>	DnaK co-chaperone; curved DNA-binding protein/uncharacterized protein	1.30 ± 0.01	1.04 ± 0.01
<i>dusB</i>	methyltransferase for 50S ribosomal subunit protein L11/tRNA-dihydrouridine synthase B	1.46 ± 0.02	1.23 ± 0.01
Membrane protein transport Processing			
<i>gsiA</i>	glutathione ABC transporter ATPase	1.09 ± 0.01	1.05 ± 0.01
<i>nlpA</i>	cytoplasmic membrane lipoprotein-28	1.64 ± 0.02	1.03 ± 0.02
<i>sufBCDS</i>	SufBCD Fe-S cluster assembly protein	1.48 ± 0.02	1.32 ± 0.01
<i>opgB</i>	OPG periplasmic biosynthetic phosphoglycerol transferases I	1.10 ± 0.01	1.06 ± 0.01
<i>setC</i>	putative arabinose efflux transporter	1.14 ± 0.02	1.10 ± 0.01
<i>oppF</i>	oligopeptide ABC transporter ATPase	1.14 ± 0.01	1.13 ± 0.01
<i>ypjK</i>	CP4-57 prophage; putative inner membrane protein	1.16 ± 0.01	1.10 ± 0.02
<i>btuB</i>	vitamin B12/cobalamin outer membrane transporter	1.05 ± 0.02	1.11 ± 0.01
<i>pitA</i>	phosphate transporter, low-affinity; tellurite importer	1.15 ± 0.01	1.17 ± 0.01
<i>cusA</i>	copper/silver efflux system, membrane component	1.15 ± 0.01	1.05 ± 0.01
Cellular Processes			
<i>yehA</i>	putative fimbrial-like adhesin protein	1.48 ± 0.03	1.24 ± 0.02
<i>motB</i>	protein that enables flagellar motor rotation	1.17 ± 0.02	1.16 ± 0.01
<i>flgFGH</i>	flagellar protein	1.56 ± 0.03	1.41 ± 0.03
<i>ypjA</i>	adhesin-like autotransporter	1.12 ± 0.02	1.07 ± 0.01
<i>tabA</i>	transcriptional repressor for divergent <i>bdCA</i> /biofilm modulator regulated by toxins; DUF386 family protein, cupin superfamily protein	1.05 ± 0.01	1.05 ± 0.01
Stress response			
<i>hslU</i>	molecular chaperone and ATPase component of HslUV protease	1.53 ± 0.03	1.29 ± 0.02
<i>flgFGH-sufBCDS- dusB-rpoA-yehA-hslU</i>		3.55 ± 0.03	1.71 ± 0.02

^a The ratio of OD₆₀₀ with CRISPRa and without.^b The ratio of pinene production with CRISPRa and without.-: means no change or negative effect.**Table 4**Effect of CRISPR interference of the down-regulated genes on growth and production of pinene in *E. coli* BW25113(P_{T5-dxs}).

Gene	Protein	Ratio of growth ^a	Ratio of pinene concentration ^b
Lipid metabolism			
<i>ydiJ</i>	putative FAD-linked oxidoreductase	1.07 ± 0.01	1.06 ± 0.02
Metabolism of cofactors and vitamins			
<i>yjbQ</i>	thiamine phosphate synthase	1.08 ± 0.02	1.08 ± 0.02
Genetic Information Processing			
<i>prpR</i>	propionate catabolism operon regulatory protein	1.06 ± 0.01	1.07 ± 0.01
<i>marR</i>	transcriptional repressor of multiple antibiotic resistance	1.22 ± 0.02	1.09 ± 0.01
<i>fabR</i>	transcriptional repressor of <i>fabA</i> and <i>fabB</i>	1.09 ± 0.01	1.07 ± 0.02
Cellular Processes			
<i>cedA</i>	cell division modulator/catalase HP11, heme d-containing	1.07 ± 0.01	1.07 ± 0.02
<i>ydiJ-yjbQ-prpR-marR-fabR-cedA</i>		1.26 ± 0.02	1.13 ± 0.02

^a The ratio of OD₆₀₀ with CRISPRi and without.^b The ratio of pinene production with CRISPRi and without.-: means no change or negative effect.

were analyzed and compared to the control strain harboring the empty vector pTargetA. The results are presented in [Table 3](#) and [Supplementary Table 3](#). The activations of 23 genes resulted in increases both in tolerance to pinene and pinene production ([Table 3](#)). The activations of 9 genes only improved pinene production. The activations of 20 genes only enhanced the tolerance to pinene.

As shown in [Table 3](#), the activations of *dxr*, *dxs*, *ispH* and *ispU* improved the production of pinene. The *dxr*, *dxs*, *ispH* and *ispU* is involved in 1-deoxy-D-xylulose-5-phosphate (DXP) pathway. These activations may increase the availability of pinene precursor isopentenyl diphosphate (IPP), enhancing the production of pinene. The activation of *ispH* improved the tolerance to pinene and the production of pinene

([Table 3](#)). The *ispH* gene encodes 4-hydroxy-3-methylbut-2-enyl diphosphate reductase which catalyzes the formation of IPP. Li et al. reported that the activation of *ispH* eliminated the accumulation of the intermediate of 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate of DXP pathway, improving the production of carotene [18]. It was reported that the mutations of Dxs^{E289G/N455I/E598V} and Dxr^{K37N/K217N} resulted in a 60% increase in the production of isoprene in *E. coli* [19]. The upregulations of the DXP pathway genes may be related to that *E. coli* YZFP is tolerant to fosfomycin [8,20]. Fosfomycin is the DXP pathway inhibitor that inhibits Dxr and methylerythritol phosphate cytidyltransferase (IspD). The Fosfomycin resistant mutants show higher levels of Dxr and IspD [21].

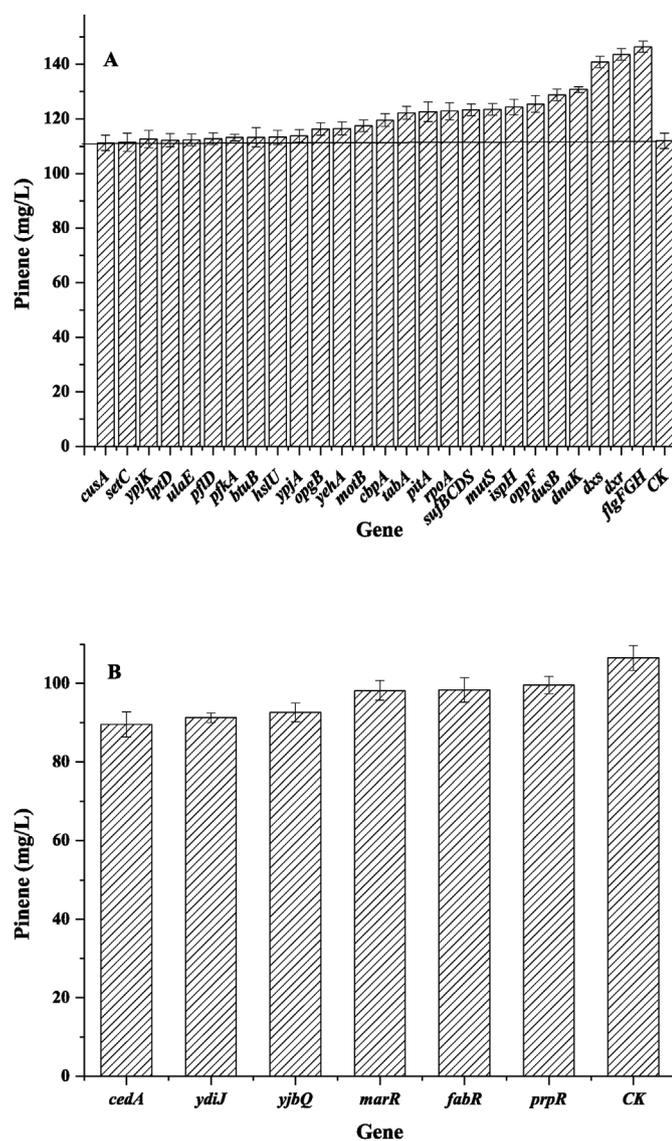


Fig. 3. Pinene production in the modular co-culture system of the whole-cell biocatalysis. (A) Using *E. coli* PINE (pZEABP-gene) and *E. coli* MEVI (pZEABP). (B) Using *E. coli* PINE (pBbB2K-dCas9*-MCPsoxS, pTargetA-X) and *E. coli* MEVI (pBbB2K-GFP, pTargetA). The data represent the means of three replicates, and error bars represent standard deviations.

Activation of the *rpoA* gene improved the tolerance to pinene and the production of pinene (Table 3). *rpoA* encodes RNA polymerase subunit alpha. The plasmid expression of the *rpoA* mutant has been successfully used to improve a lot of cellular traits of *E. coli*, including butanol tolerance [22] and the production of tyrosine [23]. This group developed the global transcription machinery engineering (gTME) to allow for global perturbations of transcriptome. Plasmid expression of the mutant *rpoD* or *rpoS* enhanced ethanol tolerance, lycopene overproduction [24] and the production of hyaluronic acid biosynthesis [25].

Activation of some membrane protein genes, such as *gsiA*, *nlpA*, *sufBCDS*, *opgB*, *setC*, *oppF*, *ypjK*, *btuB*, *pitA* and *cusA*, also improved the tolerance to pinene and the production of pinene (Table 3). These activations may be beneficial to the transport of metabolite, thereby improving tolerance and pinene synthesis. Many papers demonstrated that the membrane modified transport engineering is a powerful tool for improving tolerance and product biosynthesis [26,27].

All of 20 down-regulated genes were selected to be repressed with CRISPRi. And their effects of the repressions on growth and pinene

production were analyzed. The results are presented in Table 4 and Supplementary Table 4. Repressions of 6 genes led to increases both in tolerance to pinene and pinene production (Table 4). Repressions of 3 genes only improved the tolerance to pinene (Supplementary Table 4).

Because the phenotype in *E. coli* YZFP indeed resulted from the full mutations of these mutant genes, we tested the effect of multiple genes. However, constructing the scRNA expressing vector for simultaneous activation or repression of all genes is rather time-consuming. Because the repressions of six genes improved the tolerance and titer, we assembled the scRNA expressing vector of six genes and investigated the simultaneous effect of them. For activation, we randomly selected six genes. The simultaneous activation of the *flgFGH*, *sufBCDS*, *dusB*, *rpoA*, *yehA* and *hslU* further improved growth and pinene production than single activation (Table 3). After simultaneous activation of the six genes, the OD₆₀₀ at 12 h was increased to 0.451 ± 0.013 from 0.127 ± 0.009 without activation (Suppl. Fig. 2). From the grow curve (Suppl. Fig. 2), we can see that the wild-type strain without activation cannot grow in the present of 0.5% pinene. After activation with the six genes, the wild-type strain can grow slightly. And the tolerant strain *E. coli* YZFP can grow well in the present of 0.5% pinene.

The same result of simultaneous repression of the *ydiJ*, *yjbQ*, *prpR*, *marR*, *fabR* and *cedA* was also observed (Table 4). These results further demonstrated that the phenotype of *E. coli* YZFP is caused by the mutant genes.

3.4. Omics-guided genetic modifications

The ultimate goal of systems metabolic engineering is to further improve the production of metabolite. The above results show that the activations of 26 up-regulated mutant genes improved the tolerance to pinene and pinene production. Thus, these genes were amplified from *E. coli* YZFP and then individually cloned into plasmid pZEABP to obtain pZEABP-gene. These plasmids were individually introduced into *E. coli* PINE. These strains *E. coli* PINE (pZEABP-gene) were individually co-cultured with *E. coli* MEVI (pZEABP) with whole-cell biocatalysis.

As shown in Fig. 3A, overexpression of the mutated *cbpA*, *tabA*, *pitA*, *rpoA*, *sufBCDS*, *mutS*, *ispH*, *oppF*, *dusB*, *dnaK*, *dxs*, *dxr* and *flgFGH* increased the production of pinene by 7%, 9%, 9%, 10%, 10%, 10%, 11%, 12%, 15%, 17%, 26%, 28% and 31%, respectively. Except for *ispH*, *dxs* and *dxr*, other genes cannot be identified by general rational analysis. Some target genes difficultly found by rational analysis were identified.

Table 4 shows that the repressions of 6 down-regulated mutant genes improved the tolerance to pinene and pinene production. Thus, we also assayed whether these repressions can improve the production of pinene in the modular co-culture system of the whole-cell biocatalysis. It can be seen from Fig. 3B that these repressions in the strain tolerant to pinene did not further improve the production of pinene.

In this study, the plasmid system was used for gene expression. However, the plasmid expression system has the drawbacks of genetic instability and antibiotics pollution problem. How about chromosome expression? In addition, the simultaneous activation of the multiple genes improved the production of pinene in *E. coli* BW25113(P_{T5-dxs}) (Table 3). Are there synergies effects between these genes in the modular co-culture system of the whole-cell biocatalysis? All these should be needed further to investigate.

4. Conclusions

In this study, we combined comparative genomics and transcriptional level analysis with CRISPRa and CRISPRi to investigate the regulation mechanism of *E. coli* YZFP with higher tolerance to pinene and pinene titer. The reasons for the increased tolerance to pinene and pinene production may be associated with: 1) the mutations of the DXP pathway genes (*dxr*, *dxs*, *ispH* and *ispU*), the *rpoA* and some membrane protein genes (*gsiA*, *nlpA*, *sufBCDS*, *opgB*, *setC*, *oppF*, *ypjK*, *btuB*, *pitA* and *cusA*), and their upregulation of transcription levels; and 2) the

mutations of some genes (*ydiJ*, *yjbQ*, *prpR*, *marR*, *fabR* and *cedA*) and their downregulation of transcriptional levels. Overexpression of the mutated *cbpA*, *tabA*, *pitA*, *rpoA*, *sufBCDS*, *mutS*, *ispH*, *oppF*, *dusB*, *dnaK*, *dxs*, *dxr* and *flgFGH* genes further improved pinene production in the *E. coli*-*E. coli* modular co-culture system of whole-cell biocatalysis. This study also demonstrated that combining comparative omics analysis with CRISPRa and CRISPRi is an efficient technology to quickly find a new metabolic engineering strategy.

Acknowledgment

We are grateful to the National Natural Science Foundation of China (Grant NO. 201808248), the Natural Science Foundation of Guangdong Province (NO. 2018A030310255) and the Project of the Scientific and Technical Program of Guangzhou (No. 201607010028) for their financial support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.synbio.2019.05.001>.

References

- [1] Gomes-Carneiro MR, Viana MES, Felzenszwalb I, Paumgarten FJR, Toxicology C. Evaluation of β -myrcene, α -terpinene and (+)- and (-)- α -pinene in the Salmonella/microsome assay. *Food Chem Toxicol* 2005;43(2):247–52.
- [2] Xiang J, Luo ZH. Study on the pinene isomerization catalyzed by TiM. *Chin J Chem Eng* 2018;26(12):2537–41.
- [3] Zhao Y, Chen R, Wang Y, Yang Y. α -Pinene inhibits human prostate cancer growth in a mouse xenograft model. *Chemotherapy* 2017;63(1):1–7.
- [4] Yang JM, Nie QJ, Ren M, Feng HR, Jiang XL, Zheng YN, Liu M, Zhang HB, Xian M. Metabolic engineering of *Escherichia coli* for the biosynthesis of alpha-pinene. *Biotechnol Biofuels* 2013;6(1):60.
- [5] Sarria S, Wong B, Garcia Martin H, Keasling JD, Peralta-Yahya P. Microbial synthesis of pinene. *ACS Synth Biol* 2014;3(7):466–75.
- [6] Tashiro M, Kiyota H, Kawai-Noma S, Saito K, Ikeuchi M, Iijima Y, Umeno D. Bacterial production of pinene by a laboratory-evolved pinene-synthase. *ACS Synth Biol* 2016;5(9):1011–20.
- [7] Dunlop MJ, Dossani ZY, Szmidi HL, Chu HC, Lee TS, Keasling JD, Hadi MZ, Mukhopadhyay A. Engineering microbial biofuel tolerance and export using efflux pumps. *Mol Syst Biol* 2014;7(1):487.
- [8] Niu FX, He X, Wu YQ, Liu JZ. Enhancing production of pinene in *Escherichia coli* by using a combination of tolerance, evolution and modular co-culture engineering. *Front Microbiol* 2018;9:1623.
- [9] Kang MK, Eom JH, Kim Y, Um Y, Woo HM. Biosynthesis of pinene from glucose using metabolically-engineered *Corynebacterium glutamicum*. *Biotechnol Lett* 2014;36(10):2069–77.
- [10] Niu FX, Lu Q, Bu YF, Liu JZ. Metabolic engineering for the microbial production of isoprenoids: carotenoids and isoprenoid-based biofuels. *Synth Syst Biotechnol* 2017;2(3):167–75.
- [11] Cui YY, Ling C, Zhang YY, Huang J, Liu JZ. Production of shikimic acid from *Escherichia coli* through chemically inducible chromosomal evolution and cofactor metabolic engineering. *Microb Cell Factories* 2014;13(1):21.
- [12] Li Z, Liu JZ. Transcriptomic changes in response to putrescine production in metabolically engineered *Corynebacterium glutamicum*. *Front Microbiol* 2017;8:1987.
- [13] Li Z, Shen YP, Jiang XL, Feng LS, Liu JZ. Metabolic evolution and a comparative omics analysis of *Corynebacterium glutamicum* for putrescine production. *J Ind Microbiol Biotechnol* 2018;45(2):123–39.
- [14] Lee TS, Krupa RA, Zhang F, Hajimorad M, Holtz WJ, Prasad N, Lee S, Keasling JD. BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J Biol Eng* 2011;5(1):12.
- [15] Li XR, Tian GQ, Shen HJ, Liu JZ. Metabolic engineering of *Escherichia coli* to produce zeaxanthin. *J Ind Microbiol Biotechnol* 2015;42(4):627–36.
- [16] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 2001;25(4):402–8.
- [17] Dong C, Fontana J, Patel A, Carothers JM, Zalatan JG. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* 2018;9(1):2489.
- [18] Li YF, Lin ZQ, Huang C, Zhang Y, Wang ZW, Tang YJ, Chen T, Zhao XM. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. *Metab Eng* 2015;31:13–21.
- [19] Lv X, Gu J, Wang F, Xie W, Liu M, Ye L, Yu H. Combinatorial pathway optimization in *Escherichia coli* by directed co-evolution of rate-limiting enzymes and modular pathway engineering. *Biotechnol Bioeng* 2016;113(12):2661–9.
- [20] Kang A, Meadows CW, Canu N, Keasling JD, Lee TS. High-throughput enzyme screening platform for the IPP-bypass mevalonate pathway for isopentenol production. *Metab Eng* 2017;41:125–34.
- [21] Zhang BC, Watts KM, Hodge D, Kemp LM, Hunstad DA, Hicks LM, Odoms AR. A second target of the antimalarial and antibacterial agent fosmidomycin revealed by cellular metabolic profiling. *Biochemistry-US* 2011;50(17):3570–7.
- [22] Daniel KM, Santos CNS, Yu HM, Gregory S. Mutagenesis of the bacterial RNA polymerase alpha subunit for improvement of complex phenotypes. *Appl Environ Microbiol* 2009;75(9):2705–11.
- [23] Santos CNS, Xiao WH, Gregory S. Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in *Escherichia coli*. *Proc Natl Acad Sci U S A* 2012;109(34):13538–43.
- [24] Alper H, Stephanopoulos G. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng* 2007;9(3):258–67.
- [25] Yu H, Tyo K, Alper H, Klein-Marcuschamer D, Stephanopoulos G. A high-throughput screen for hyaluronic acid accumulation in recombinant *Escherichia coli* transformed by libraries of engineered sigma factors. *Biotechnol Bioeng* 2010;101(4):788–96.
- [26] Dunlop MJ. Engineering microbes for tolerance to next-generation biofuels. *Biotechnol Biofuels* 2011;4(1): 32–32.
- [27] Qi YL, Liu H, Chen XL, Li LM. Engineering microbial membranes to increase stress tolerance of industrial strains. *Metab Eng* 2019;53:24–34.