UC Berkeley
UC Berkeley Previously Published Works

Title

Lepidopteran mevalonate pathway optimization in Escherichia coli efficiently produces isoprenol analogs for next-generation biofuels

Permalink

Authors

Pang, Bo
Li, Jia
Eiben, Christopher B
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Lepidopteran mevalonate pathway optimization in *Escherichia coli* efficiently produces isoprenol analogs for next-generation biofuels

Bo Pang[a,d,e,#], Jia Li[a,b,d,#], Christopher B Eiben[a,d], Ethan Oksen[c,d], Carolina Barcelos[c,d], Rong Chen[a,d,f], Elias Englund[a,d], Eric Sundstrom[c,d], Jay D Keasling[a,d,e,g,h,*]

[a] Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA 94608, United States

[b] State Key Laboratory of Biocatalysis and Enzyme Engineering, Environmental Microbial Technology Center of Hubei Province, School of Life Sciences, Hubei University, Wuhan, Hubei 430062, P. R. China

[c] Advanced Biofuels & Bioproducts Process Development Unit, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[d] Biological Systems and Engineering, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States

[e] Department of Chemical & Biomolecular Engineering, University of California, Berkeley, CA 94720, United States

[f] School of Public Health, Hangzhou Normal University, Hangzhou, Zhejiang 311121, P. R. China

[g] Novo Nordisk Foundation Center for Biosustainability, Technical University Denmark, DK 2970 Horsholm, Denmark

[h] Center for Synthetic Biochemistry, Shenzhen Institutes for Advanced Technologies, Shenzhen, Guangdong 518055, P. R. China

[#] These authors contributed equally: Bo Pang and Jia Li

[*] Corresponding author:

Abstract

Terpenes constitute the largest class of natural products with over 55,000 compounds with versatile applications including drugs and biofuels. Introducing structural modifications to terpenes through metabolic engineering is an efficient and sustainable way to improve their properties. Here, we report the optimization of the lepidopteran mevalonate (LMVA) pathway towards the efficient production of isopentenyl pyrophosphate (IPP) analogs as terpene precursors. First, we linked the LMVA pathway to NudB, a promiscuous phosphatase, resulting in the production of the six-carbon analog of 3-methyl-3-buten-1-ol (isoprenol), 3-ethyl-3-buten-1-ol (C6-isoprenol). Using C6-isoprenol as the final product, we then engineered the LMVA pathway by redirecting its upstream portion from a thiolase-dependent pathway to a beta-oxidation pathway. The beta-oxidation LMVA pathway transforms valeric acid, a platform chemical that can be produced from biomass, into C6-isoprenol at a titer of 110.3 mg/L, improved from 5.5 mg/L by the thiolase LMVA pathway, which used propionic acid as a feedstock. Knockout of the *E. coli* endogenous thiolase genes further improved the C6-isoprenol titer to 390 mg/L, implying efficient production of homo isopentenyl pyrophosphate (HIPP). The beta-oxidation LMVA-NudB pathway also converts butanoic acid and hexanoic acid into isoprenol and isoprenol's seven-carbon analog, 3-propyl-3-buten-1-ol (C7-isoprenol), respectively, suggesting the beta-oxidation LMVA pathway produces IPP and C7-IPP from the corresponding fatty acids. Fuel property tests revealed the longer chain isoprenol analogs have lower water solubilities, similar or higher energy densities, and comparable research octane number (RON) boosting effects to isopentenols. This work not only optimizes the LMVA pathway, setting the basis for homoterpene biosynthesis to expand terpene chemical space, but provides an efficient pathway to produce isoprenol analogs as next-generation biofuels from sustainable feedstocks.

Keywords

Lepidopteran mevalonate pathway

Thiolase

Beta oxidation

Homo isopentenyl pyrophosphate

Biofuel

1. Introduction

Terpenes are natural products with diverse functions in different organisms. These compounds have benefited human society since antiquity, applied as materials (e.g., natural rubber) (Men et al., 2018), traditional medicines (e.g., crude extracts of *Artemisia annua*) (Tu, 2011), pharmaceuticals (e.g., artemisinin and taxol) (Cox-Georgian et al., 2019), and cosmetics (e.g., squalene) (Huang et al., 2009). The versatile functions and applications of terpenes owe to their diverse chemical structures, with more than 55,000 molecules known (Ajikumar et al., 2008). Nature employs a concise paradigm to build such structures, in which isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), five-carbon units derived from the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway (Rohmer, 1999) or mevalonate (MVA) pathway (Buhaescu and Izzedine, 2007), are first polymerized head-to-tail into prenyl pyrophosphates, including geranyl (GPP, 10 carbons), farnesyl (FPP, 15 carbons), and geranylgeranyl (GGPP, 20 carbons) pyrophosphates. Then, the prenyl pyrophosphates are processed, mostly cyclized, by terpene synthases to produce terpene scaffolds (Oldfield and Lin, 2012). Lastly, the optional tailoring proteins, including P450s (Zheng et al., 2019) and glycosyltransferases (Rivas et al., 2013), modify the scaffolds to afford the final products. Through this paradigm, structural diversities of the final products are introduced from different numbers of five-carbon units polymerized, varied cyclizations (different scaffolds), and optional post modifications. The final products are usually with multiple of five carbons. Terpenes with non-multiple of five carbons are rare in nature and challenging to sample, even though this chemical space is valuable to explore for novel applications.

We reported heterologous expression of the lepidopteran mevalonate (LMVA) pathway, a propionyl-CoA ligase, and terpene cyclases in *E. coli* to produce several novel sesquiterpene analogs containing 16 carbons (Eiben et al., 2019). The LMVA pathway produces 6-carbon analogs of IPP and DMAPP, homo-IPP (HIPP) and homo-DMAPP (HDMAPP), in a reaction sequence highly similar to the canonical MVA pathway (Schooley et al., 1973). The difference is that the thiolase condenses a propionyl-CoA and an acetyl-CoA making 3-ketovaleryl-CoA in the LMVA pathway instead of the condensing two acetyl-CoA

to produce acetoacetyl-CoA in the canonical MVA pathway (**Fig. 1**). The "extra" carbon from the propionyl-CoA is transformed into HIPP/HDMAPP, combined with two 5-carbon isoprenoid precursors to afford the sixteen-carbon (C16) products. With the LMVA pathway expressed in *E. coli*, we established a biosynthesis platform for novel homoterpenes deviating from the "multiple of five carbons" rule.

While the previous study successfully produced the final homosesquiterpenes, the details of the LMVA pathway, especially the production of HIPP and HDMAPP in the *E. coli* host was not assayed. Moreover, the production needs optimization because low C16 terpene titers have hindered us from accumulating and purifying these products for characterization. Here, we investigate the LMVA pathway by introducing a promiscuous phosphatase, NudB (Chou and Keasling, 2012; Zada et al., 2018), to hydrolyze the terpene precursors to their corresponding alcohols. These alcohols are readily detected by gas chromatography (GC). Using the alcohols as the final product, we were able to engineer and optimize the LMVA pathway to increase the production of HIPP, the direct product of LMVA and the starting substrate for homoterpenes synthesis. Also, the higher 3-methyl-3-buten-1-ol (isoprenol) analogs have excellent fuel properties, making them candidates for next-generation biofuels.
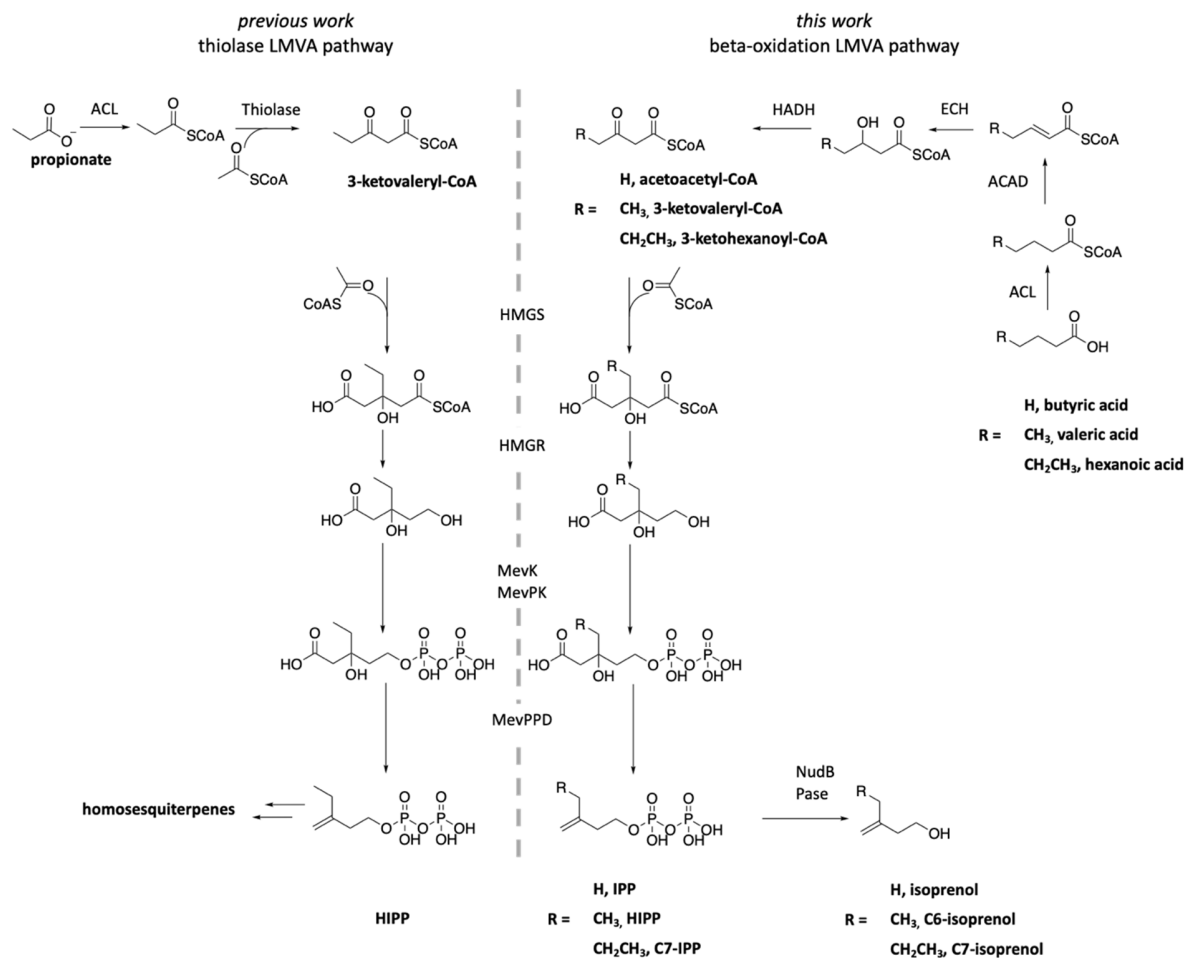
Fig. 1 Comparison of the lepidopteran mevalonate (LMVA) pathways leading to higher isopentenyl pyrophosphate analogs in *E. coli*. The previous pathway (Eiben et al., 2019) depends on a thiolase to produce the key intermediate, 3-ketovaleryl-CoA, while the pathway here employs beta-oxidation enzymes to afford this intermediate and the analogs thereof. Acronyms: ACL, CoA ligase; ACAD, acyl-CoA dehydrogenase; ECH, enoyl-CoA hydratase; HADH, 3-hydroxyacyl-CoA dehydrogenase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MevK, mevalonate kinase; MevPK, mevalonate phosphate kinase; MevPPD, mevalonate pyrophosphate decarboxylase; NudB, pyrophosphatase; Pase, phosphatase; IPP, isopentenyl pyrophosphate; HIPP, homo isopentenyl pyrophosphate; C7-IPP, 3-propyl-3-buten-1-yl pyrophosphate.

Isoprenol, 3-methyl-3-buten-1-ol; C6-isoprenol, 3-ethyl-3-buten-1-ol; C7-isoprenol, 3-propyl-3-buten-1-ol

2. Materials and methods

2.1 General materials and methods

For *E. coli* growth, pre-poured Luria-Bertani (LB) agar plates (Teknova Inc, CA) and LB broth (miller, Merck KGaA, Darmstadt, Germany) were used. For *E. coli* production runs, Terrific Broth (Yeast Extract 24g, Tryptone 12g, K_2HPO_4 9.4g, KH_2PO_4 2.2g, pH = 7.2, Molecular Biologicals International, Inc., CA) was used. Antibiotics were added in these media when needed at the following concentrations: carbenicillin, 100 mg/L, chloramphenicol, 25 mg/L. 3-ethyl-3-buten-1-ol (C6-isoprenol), (Z)-3-methyl-2-penten-1-ol ((Z)-C6-prenol) and 3-propyl-3-buten-1-ol (C7-isoprenol) were purchased from Enamine Ltd, NJ. Other chemicals were purchased from Millipore Sigma (Merck KgaA, Darmstadt, Germany) unless otherwise indicated. Bacterial strains were listed in **Table S1**.

2.2 Plasmid construction

All the plasmids and primers were designed using j5 DNA assembly (Hillson et al., 2012) unless otherwise indicated, and listed in **Table S1** and **Table S2**, respectively. The plasmids are publicly available through the Joint BioEnergy Institute (JBEI) registry (https://public-registry.jbei.org/folders/693). Primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA). PCR amplifications were performed on an Applied Biosystems Veriti Thermal Cycler (Thermo Fisher Scientific, MA) using Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific, MA) or PrimeSTAR Max DNA Polymerase (Takara Bio Inc., Japan). Gene codon optimization was conducted using the IDT Codon Optimization Tool (IDT, Coralville, IA). Gene fragments were synthesized by Integrated DNA Technologies (IDT, Coralville, IA). Plasmid isolation was carried out using QIAprep Spin Miniprep Kit (Qiagen, Germany) or by the plasmid DNA preparation service

provided by Genewiz, South San Francisco, CA. DNA gel extractions were performed using Zymoclean

Gel DNA recovery Kit (ZYMO Research, Irvine, CA). Sanger sequencing of plasmids and bacterial

clones was supplied by Genewiz, South San Francisco, CA. The plasmid sequences were verified using

the DIVA sequence validation service performed by the Synthetic Biology Informatics Group of the Joint

BioEnergy Institute (Chiniquy et al., 2020; Thompson et al., 2018). DNA gel photos were taken using

BioSpectrum Imaging System (Cambridge, UK).

2.2.1 Plasmids for beta-oxidation lepidopteran mevalonate pathway

pJH53 and pJH55 were constructed using Golden Gate Assembly using the BsaI, as previously described

(Eiben et al., 2019).

To construct pJH53, the following three fragments were amplified and assembled: a 5.0-kb fragment

amplified from pJH10 using primer pair *pJH53_pJH10_F1/pJH53_pJH10_R1*; a 5.5-kb fragment

amplified from pJH10 using primer pair *pJH53_pJH10_F2/pJH53_pJH10_R2*; a 2.2-kb fragment,

containing gene encoding CsAAE1, amplified from pSH14 using primer pair

pJH53_pSH14_F/pJH53_pSH14_R.

To construct pJH55, the following three fragments were amplified and assembled: a 1.1-kb fragment

amplified from pJH15 using primer pair *pJH55_pJH15_F1/pJH55_pJH15_R1*; a 6.3-kb fragment

amplified from pJH15 using primer pair *pJH55_pJH15_F2/pJH55_pJH15_R2*; a 4.6-kb fragment,

containing genes encoding FadB and acyl-CoA dehydrogenase, amplified from pEG1780 using primer

pair *pJH55_pEG1780_F/ pJH55_pEG1780_R.*

2.2.2 Plasmids for isoprenol analogs production

pJL01 and pJL02 derivatives were constructed using the NEBuilder HiFi DNA Assembly (New England

Biolabs, MA).

pJL01-MlD was constructed by replacing the Fpps1, Fpps2, and epi-Isozizaene synthase coding genes in pJH55 by the NudB coding gene in pNudB. For construction, the following three fragments were amplified and then assembled: a 4.2-kb fragment amplified from pJH55 using primer pair *pJL01-MlD_pJH55_F1/pJL01-MlD_pJH55_R1*; a 4.1-kb fragment amplified from pJH55 using primer pair *pJL01-MlD_pJH55_F2/pJL01-MlD_pJH55_R2*; a 0.5-kb fragment amplified from pNudB using primer pair *pJL01-MlD_pNudB_F/pJL01-MlD_pNudB_R.*

To substitute the acyl-CoA dehydrogenase coding gene in pJL01-MlD with homologs, a 6.5-kb fragment was first amplified from pJL01-MlD using primer pair *pJL01_F/pJL01_R*. The resulting fragment was assembled with the synthesized DNA fragments, *MeD* and *PpD* listed in **Table S3**, to yield the plasmids pJL01-MeD and pJL01-PpD, respectively.

pJL02-CsCL was constructed by deleting the IDI coding gene in pJH53. For construction, the following two fragments were amplified from pJH53 and assembled: a 6.2-kb fragment using primer pair *pJL02-CsCL_pJH53_F1/pJL02-CsCL_pJH53_R1*; a 5.7-kb fragment using primer pair *pJL02-CsCL_pJH53_F2/pJL02-CsCL_pJH53_R2.*

To substitute the hexanoyl-CoA ligase coding gene in pJL02-CsCL with homologs, the following two fragments were first amplified from pJL01-MlD: a 5.0-kb fragment using primer pair *pJL02_F1/pJL02_R1*; a 4.7-kb fragment using primer pair *pJL02_F2/pJL02_R2*. The resulting fragments were assembled with the synthesized DNA fragments, *ScPCL*, *PcCL*, *SsCL*, and *SaCL* listed in **Table S3**, to yield the plasmids pJL02-ScPCL, pJL02-PcCL, pJL02-SsCL, and pJL02-SaCl, respectively.

2.3 Gene knockout in *E. coli*

The *atoB* knockout and *atoB yqeF* double knockout strains were generated using the λ red recombinase protocol previously described (Baba et al., 2006). A kanamycin resistance cassette flanked by FLP recognition target sites was amplified from the plasmid pKD13 with primers *KO_kan_F* and *KO_kan_R*. The 1000 base pair homology regions upstream and downstream of *atoB* and *yqeF* were amplified using the primer pairs *atoB_US_F/atoB_US_R* and *atoB_DS_F/atoB_DS_R*, and *yqeF_US_F/yqeF_US_R* and *yqeF_DS_F/yqeF_DS_R* using BW25133 genomic DNA as template, respectively. The homology regions were designed to knock out the entire gene except for the starting ATG and the last 21 base pairs to avoid disrupting potential downstream ribosomal binding sites. The upstream and downstream homology sequences were combined with the kanamycin cassette by using NEBuilder HiFi DNA Assembly (New England Biolabs), and a second PCR was done to amplify the three-part fusion DNA sequences using primer pairs *atoB_US_F/atoB_DS_R* or *yqeF_US_F/yqeF_DS_R* using the corresponding assembly product.

To generate marker-free *atoB* knockout, *E. coli* 6C01, the pKD46 plasmid with the λ red recombination genes and a temperature-sensitive replicon was transformed into chemically competent *E. coli* BL21(DE3) cells and selected on carbenicillin LB agar plates at 30°C. Plasmid-containing cells were made electrocompetent, and expression of recombination genes was induced with 0.1% arabinose. ~600 ng of the three-part PCR product was introduced into the cells by electroporation, and positive colonies were selected for on LB kanamycin plates grown at 37°C overnight. To remove the kanamycin resistance cassette, pCP20 was transformed into the selected kanamycin resistance cells using electroporation and selected on carbenicillin and kanamycin LB agar plates at 30°C. The resulting colonies were streaked on an LB plate without antibiotic and grew overnight at 42°C to remove the cassette. The resulting colonies were streaked on an LB plate without antibiotic, a carbenicillin LB agar plate, and a kanamycin LB agar plate, respectively, to confirm the loss of the kanamycin cassette and pCP20. Finally, the knockout of *atoB* was confirmed via colony PCR using the primer pair *atoB_C_F/atoB_C_R*. (**Fig. S2**)

To generate maker-free *atoB yqeF* double knockout, *E. coli* 6C02, the same protocol for *atoB* knockout was employed. The three-part PCR product was introduced into 6C01. The knockout of *yqeF* was confirmed via colony PCR using the primer pair *yqeF_C_F/yqeF_C_R.* (**Fig. S2**)

2.4 *E. coli* production runs

Two plasmids bearing the pathway genes were co-transformed into *E. coli* expression hosts using a room temperature electroporation method, in which the electrocompetent bacterial cells were prepared at room temperature freshly (Tu et al., 2016). The transformed cells were plated on LB agar plates added with carbenicillin and chloramphenicol, and the resulting plates were incubated at 30°C until colonies appeared. The freshly transformed single colonies were inoculated into 2 mL production media (TB medium containing 0.4% glycerol (w/v), 0.4% glucose (w/v), carbenicillin, and chloramphenicol, pH = 7.2) in 5 mL BD falcon test tubes, incubated at 30°C, 200 rpm for 16 h. The grown cultures were inoculated (1:100 v/v) into 10 mL production media in 25 mL glass tubes, and the cultures were grown at 37°C, 200 rpm until OD_{600} reached 0.8. After incubating on ice for 10 min, isopropyl-β-D-thiogalactopyranoside (IPTG, 1 mM) and substrates (1 g/L valeric acid, 0.88 g/L butyric acid, 1.16 g/L hexanoic acid, or 1 g/L sodium propionate) were added to the cultures, respectively. The cultures were incubated at 18°C, 200 rpm for 72 h to produce isoprenol analogs. After the production, a ten-fold dilution of the production broth (50 μL of the production broth was mixed with 450 μL of the fresh production media) was subjected to endpoint optical density (OD600) measurement using the cuvette port of a SpectraMax M2e plate reader (Molecular Devices, CA). All the production runs were conducted in biological triplicate.

2.5 Analytic chemistry

2.5.1 Sample processing

200 μL of production broth was transferred into a 1.5-mL eppendorf tube, followed by adding 4 μL hydrochloric acid (6 mol/L). Then, 400 μL of ethyl acetate containing 30 mg/L butanol, as the internal

standard, was added. The mixture was shaken vigorously (3000 rpm) on a vortex shaker for 15 min. The resulting mixture was centrifuged at 12000 g for 1 min. Then, 300 μL of the organic layer was collected into a gas chromatography (GC) vial. To prepare standards for calibration curves, the analytic targets were added into the production media to make a series of standard solutions. The standard solutions were processed using the same method for samples.

2.5.2 GC-flame ionization detection (GC-FID) analysis

GC-FID analysis was performed on a Thermo Scientific FOCUS GC system with TriPlus AS Autosampler using a DB-WAX column (L 30 m * ID 0.25 mm, Agilent technologies, CA) with a helium flow at a constant pressure of 150 kPa (the length of the column changed because of trimming, resulting in shifting retention times in different batches). The inlet temperature was 200°C, and the base temperature of the FID was 250°C. 1 μL sample was injected using the splitless mode. The oven starting temperature was 40°C held for 3.5 min. The temperature increased at 10°C per minute until 120°C. Then the temperature was increased at 20°C per minute until 160°C. Then the temperature was increased at 100°C per minute until 245°C and held for 10 min.

2.5.2 GC-mass spectrometry (GC-MS) analysis

GC-MS analysis was performed on an Agilent Intuvo 9000 system equipped with the pneumatic switching device (PSD) using two tandem DB-WAX columns (L 15 m * ID 0.25 mm, Agilent technologies, CA) with a helium flow of 1 mL/min at the first column and 1.2 mL/min at the second column. The inlet temperature, the Intuvo flow path temperature, and the MS transfer line temperature are 250°C. 1 μL sample was injected using splitless mode. The oven starting temperature was 60°C held for 2 min. The temperature increased at 15°C per minute until 120°C. Then the temperature was increased at 30°C per minute until 245°C. After the temperature ramping program, a post-run was conducted at 250°C for 5 min with a helium flow of -2.691 mL/min at the first column and 3.106 mL/min at the second column.

2.6 Fuel properties analysis

2.6.1 Water solubility estimation using high performance liquid chromatography (HPLC)

Water solubility of isoprenol analogs was estimated using their logP, octanol-water partition coefficient, values via the following equation:

$$logSw = A - B * logP$$

where Sw = molar aqueous solubility, logP = octanol-water partition coefficient.

To calculate the values of A and B, the experimental logP and water solubility values of 1-octanol, 1-heptanol, 1-hexanol, 1-pentanol, and 1-butanol were used. A simple linear recession of logSw against logP values resulted in A = 0.9939, B = 1.103, and R^2 = 0.9955. (**Fig. S3**)

To determine logP values, a HPLC based method, referred to EPA OPPTS 830.7570, was conducted via the following equation:

$$logP = C + D * logk$$

where k is the capacity factor, given by the expression

$$k = (t_R - t_0)/t_0$$

where t_R is the retention time of the test substance, and t_0 is the dead time.

The determination of t_R and t_0 was performed on an Agilent 1260 HPLC system (Agilent Technologies Inc., CA) equipped with a refractive index detector (RID). 20 μL of the samples and alcohol standards, prepared at 100 mg/L in 50% methanol water solution were subjected to HPLC-RID analysis on an Agilent Zorbax column (SB-C18, 5 μm, 4.6 x 250 mm, Agilent Technologies Inc., CA) heated to 30°C by isocratic elution of a 40% (v/v) acetonitrile water solution. 100 mg/L thiourea aqueous solution was used to determine t_0. The retention times were measured in technical triplicate.

A simple linear recession of experimental logk against logP values of the linear primary alcohols resulted in C = 1.276, D = 1.942, and R^2 = 0.9932.

2.6.2 Energy density and research octane number (RON) analyses

The gross heat of combustion (higher heating value) was measured using the ASTM D4809 method, performed by Intertek (Boston, MA, USA). For RON analysis, the alcohols were blended with reformulated blendstock for oxygenate blending (RBOB) gasoline at 10% by volume, respectively. The resulting 10% alcohol blends were analyzed using the Advanced Fuel Ignition Delay Analyzer (AFIDA) method (Luecke and Zigler, 2021), performed at National Renewable Energy Laboratory.

2.7 Data processing

GC-FID data were collected using ChromQuest 4.2.34 (Thermo Fisher Scientific, MA) and processed using OpenChrom Lablicate Edition 1.4 (Wenig and Odermatt, 2010). GC-MS data were collected using MassHunter GC/MS Acquisition B.07.06.2704 (Agilent Technologies Inc., CA) and processed using MSD ChemStation F.01.03.2357 (Agilent Technologies Inc., CA). HPLC data were collected and processed using OpenLAB CDS ChemStation Edition A.02.10 (Agilent Technologies Inc., CA). DNA gel photos were taken and processed using VisionWorks Acquisition and analysis 8.2 (Analytik Jena US, Upland, CA). Statistics were conducted using Microsoft Excel for Mac 16.49 and Prism 8 (GraphPad, San

Diego, CA). GC-FID chromatographies, MS spectra, and histograms were exported using Prism 8 (GraphPad, San Diego, CA) using the processed data.

3. Results

3.1 Verification of the terpene units produced by the thiolase lepidopteran mevalonate pathway

The homosesquiterpene biosynthesis pathway we constructed in *E. coli* contains two sections to incorporate propionate into $C_{16}H_{26}$ terpenes. In the first section, the LMVA pathway transforms propionate into HIPP and HDMAPP. In the second section, these C6 building blocks are transformed into terpenes (**Fig. 1**). The production of those $C_{16}H_{26}$ terpenes was previously confirmed using GC-MS and GC-MS-TOF analysis and ^{13}C labeling experiments. However, several challenges hindered us from purifying the final products for detailed structural characterization via nuclear magnetic resonance (NMR): 1) the $C_{16}H_{26}$ terpenes were produced at low titers, 1.05 mg/L in the best case and 2) the ratio of C16 terpenes to total terpenes was modest, 11.25% in the best scenario. Moreover, this pathway was expected to produce higher terpenes with molecular formulas of $C_{17}H_{28}$ and $C_{18}H_{30}$, but we did not detect these terpenes. We suspected the low production and low ratio of the homoterpenes may be ascribed to insufficient production of the C6 isoprenoid precursors, HIPP and HDMAPP.

To assess the production of the C6 isoprenoid precursors, we tested the production of their corresponding alcohols in an *E. coli* strain co-expressing the LMVA pathway and NudB. NudB is a phosphatase and works with another endogenous phosphatase to dephosphorylate C5-C15 prenyl diphosphates into their corresponding alcohols in *E. coli* (Chou and Keasling, 2012; Zada et al., 2018). To construct the production stain, we transformed pNudB (Chou and Keasling, 2012) and pJH10 (Eiben et al., 2019) into *E. coli* BAP1 (Pfeifer et al., 2001). The resulting strain was grown and induced for production, with sodium propionate added at the time of induction. After production, we extracted the production broth using ethyl acetate, and the organic phase was analyzed by GC-FID and GC-MS. In the analysis, GC-FID detected a peak in the production broth, and this peak was absent in the negative control (**Fig. 2A-i, iii**).

This peak also has the same retention time as the C6-isoprenol standard (**Fig. 2A-i, v**). GC-MS indicated the mass spectrum of the new peak is similar to the spectrum of the C6-isoprenol standard, with major peaks at $m/z = 67, 55, 82$, indicating the new peak has similar electron ionization (EI) induced fragmentation to C6-isoprenol (**Fig. 2B-i, v**). GC analysis confirmed the production of C6-isoprenol by the production strain, suggesting HIPP was produced. To evaluate the production level of HIPP, we quantified the C6-isoprenol at 5.5 mg/L using the C6-isoprenol standard curve (**Fig. 2C**). In the GC analysis, we also noticed the production of isoprenol (**Fig. 2A-i, v**), the hydrolyzed product of IPP, at 4.1 mg/L (**Fig. 2C**), indicating the production strain also produces a comparable amount of IPP as HIPP. We reasoned that IPP might come from the E. *coli* endogenous MEP pathway and the promiscuity of the LMVA pathway, in which the thiolase accepted two molecules of acetyl-CoA. We also attempted to confirm the production of (Z)-C6-prenol, the proposed hydrolyzed product of HDMAPP, using a synthesized standard with GC-FID and GC-MS. However, (Z)-C6-prenol was not detected in the production broth, suggesting the *idi* gene, encoding the isopentenyl-diphosphate delta-isomerase, might not function normally to transform HIPP to HDMAPP (**Fig. 2A-i, v**). The results of these production runs indicated the supplement of the C6 isoprenoid precursors from the LMVA pathway is at a low level, comparable to the supply of the normal C5 isoprenoid precursors.
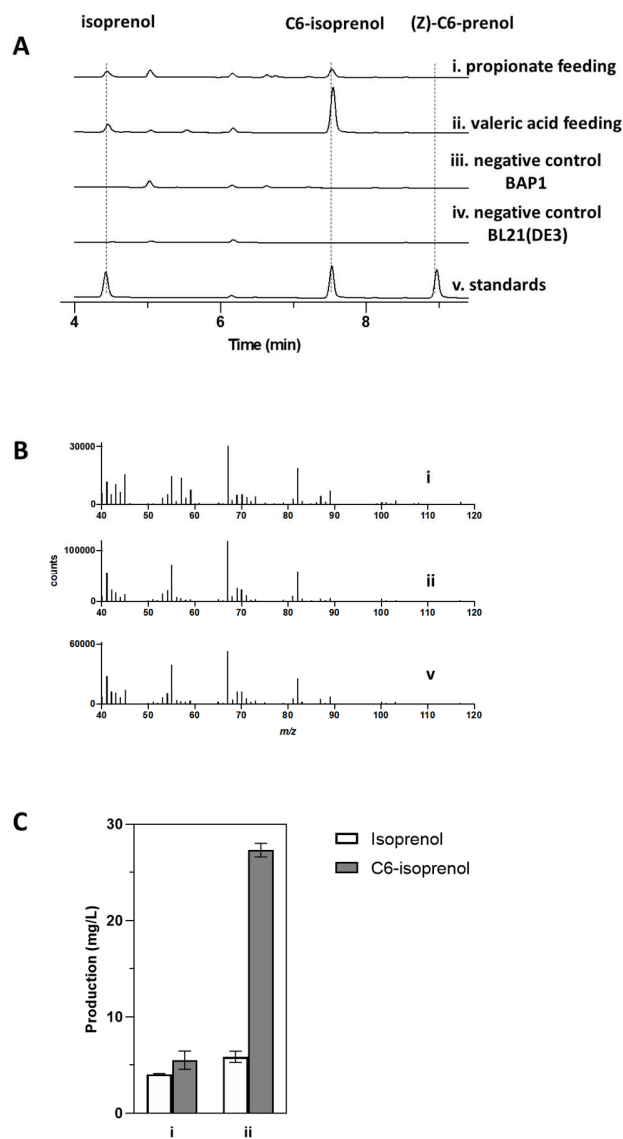
Fig. 2 Production of isoprenol and C6-isoprenol in *E. coli* strains co-expressing NudB and the thiolase LMVA pathway (i) or the beta-oxidation LMVA pathway (ii), along with the non-pathway negative controls (iii, iv). (A) GC-FID traces of the ethyl acetate extracted production broth and the strands (v) of isoprenol, C6-isoprenol, and (Z)-3-methyl-2-penten-1-ol ((Z)-6C-prenol). (B) Mass spectrums of C6-isoprenol in the production broths and standards. (C) Quantification of the isoprenol and C6-isorpenol production in the production broths.

3.2 Engineering the LMVA pathway to increasing the production of C6-isoprenol

To address the insufficient supplement of HIPP, we revisited the LMVA pathway for C6-isoprenol production in *E. coli*. Like the natural MVA pathways, the previous work employed a thiolase to condense propionyl-CoA and acetyl-CoA to afford 3-ketovaleryl-CoA (**Fig. 1**, thiolase LMVA pathway). Because the predicted thiolase from *Bombyx mori* does not express in *E. coli*, we used PhaA from *Acinetobacter* strain RA3849. PhaA has been well characterized in PHA/PHB production, in which PhaA accepts propionyl-CoA and acetyl-CoA. However, PhaA also converts two acetyl-CoA to acetoacetyl-CoA (Aldor and Keasling, 2001), an intermediate for IPP production via the thiolase LMVA pathway. Moreover, a previous study has shown that PhaA homologs catalyze the degradation reaction better than the condensation reaction (Kim and Kim, 2014), suggesting that the degradation of 3-ketovaleryl-CoA is a favored direction in the PhaA-catalyzed reversible reaction. Hence, we suspected the thiolase catalyzed reaction is the bottleneck in the thiolase LMVA pathway for HIPP production.

To bypass the thiolase in the pathway, we chose a new source for 3-ketovaleryl-CoA, namely beta-oxidation of valeric acid. Like fatty acid oxidation, the new pathway first activates valeric acid, a biomass-derived chemical, into valeryl-CoA by a CoA ligase. An acyl-CoA dehydrogenase then oxidizes valeryl-CoA into 2-pentenyl-CoA, which is then transformed by an enoyl-CoA hydratase and a 3-hydroxyacyl-CoA dehydrogenase into 3-ketovaleryl-CoA (**Fig. 1**, beta-oxidation LMVA pathway). We believed this pathway would be more efficient than the PhaA-dependent pathway for bypassing the highly reversible and promiscuous thiolase-catalyzed reaction.

We then tested the proposed beta-oxidation LMVA pathway, which accepts valeric acid into C6-isoprenol. Two plasmids were constructed to express the pathway (**Fig. S1**): 1) pJL01, modified from pJL15, containing genes encoding NudB, *Micrococcus luteus* acyl-CoA dehydrogenase, and *E. coli* FadB, a bifunctional enzyme that catalyzes the hydration and dehydrogenation reactions (Goh et al., 2012); 2) pJL02, modified from pJL10, containing the genes that encode the enzymes in the LMVA pathway and a

Cannabis sativa acyl activating enzyme, CsAAE1, for valeryl-CoA production (Luo et al., 2019; Stout et al., 2012). *E. coli* BL21(DE3) transformed with pJL01 and pJL02 was grown in the presence of valeric acid and induced to express the pathway genes. GC-FID (**Fig. 2A-ii, iv, v**) and GC-MS (**Fig. 2A-ii, v**) confirmed the production of C6-isoprenol at 27.3 mg/L (**Fig. 2C**). In the production broth, we also detected isoprenol at 5.8 mg/L (**Fig. 2A-II and Fig. 2C**). Compared to the thiolase LMVA pathway, C6-isoprenol production increased fourfold, and the ratio of C6-isoprenol with isoprenol increased to 4.7 from 1.3.

3.3 Optimization of the beta-oxidation LMVA pathway in *E. coli*

Next, we screened CoA ligase and acyl-CoA dehydrogenase homologs, catalyzing the beta-oxidation pathway's first two steps. We chose CoA ligase and acyl-CoA dehydrogenase candidates that have been overexpressed in *E. coli* and assayed in vitro in previous studies. Moreover, most of these candidates have known kinetic data for medium chain fatty acid substrates (**Table 1**). For the CoA ligase, we selected the phenylacetate-CoA ligase from *Streptomyces coelicolor* A3(2) (ScCL) (Go et al., 2012), ORF26 from *Streptomyces aizumensis* (SaCL) (Zhang et al., 2017), the medium-chain fatty acyl-CoA ligase from *Streptomyces* sp. SN-593 (SsCL) (Miyazawa et al., 2015), and the CoA ligase from *Penicillium chrysogenum* (PcCL) (Koetsier et al., 2011), in addition to CsAAE1 (CsCL) from *Cannabis sativa*. For the acyl-CoA dehydrogenase, we selected the short-chain acyl-CoA dehydrogenase from *Pseudomonas putida* KT2440 (PpD) (McMahon et al., 2005) and the butyryl-CoA dehydrogenase from *Megasphaera elsdenii* (MeD) (DuPlessis et al., 1998), in addition to the acyl-CoA oxidase (MlD) from *Micrococcus luteus*. The gene sequences were codon-optimized for *E. coli* and cloned into the pJL02 and pJL01 plasmid series (**Fig. S1, Table S3**).

With the constructs in hand, we tested C6-isoprenol production in *E. coli*. First, we screened the CoA ligases, catalyzing the first step in beta-oxidation, with the well-characterized MeD as the acyl-CoA dehydrogenase. After production, we used GC-FID to quantify C6-isoprenol in the broths. The C6-

isoprenol production varied substantially with different CoA ligases, suggesting this is a key step in the whole pathway. Among the genes we tested, PcCL gave rise to the highest production at 58.6 mg/L (**Fig. 3**), a result consistent with the reported kinetics data (**Table 1**). Also, we noticed that the combination of CsCL with MeD had a decreased production of 6.9 mg/L, down from the 27.3 mg/L produced by CsCL/MlD, suggesting that MlD is a better homolog for the second step (**Fig. 3**). This hypothesis was supported by the screening results of the second step catalyzed by acyl-CoA dehydrogenase. The PcCL and MlD pair gave the highest C6-isoprenol production at 110.3 mg/L, so we used this pair of genes for the first two steps of the beta-oxidation pathway in the following experiments. Endpoint optical density (OD600) analysis was performed to evaluate the impact of the expression of different beta-oxidation genes on cell growth. The results indicated the expression of the CoA ligases and acyl-CoA dehydrogenase barely impacts the growth. (**Fig. 3**) The slight growth inhibition effect in the high C6-isoprenol production, e.g., the PcCL/MlD, may result from the toxicity of C6-isoprenol. (Kang et al., 2019; Dunlop, 2011)

Table 1

Tested CoA ligase and acyl-CoA dehydrogenase homologs

Protein	Abbreviation	Organism	k_{cat}/K_M $(M^{-1}S^{-1})^*$	Reference
CoA Ligase (ACL)				
CsAAE1	CsCL	*Cannabis sativa*	541 Hexanoate	Stout et al. 2012
Phenylacetate-CoA ligase	ScCL	*Streptomyces coelicolor* A3(2)	110 Valerate	Go et al. 2012
ORF26	SaCL	*Streptomyces aizunensis*	16.7 Valerate	Zhang et al. 2017
Medium chain fatty acyl-CoA ligase	SsCL	*Streptomyces* sp. SN-593	272 Hexanoate	Miyazawa et al. 2015
CoA ligase	PcCL	*Penicillium*	10675	Koetsier et al. 2011

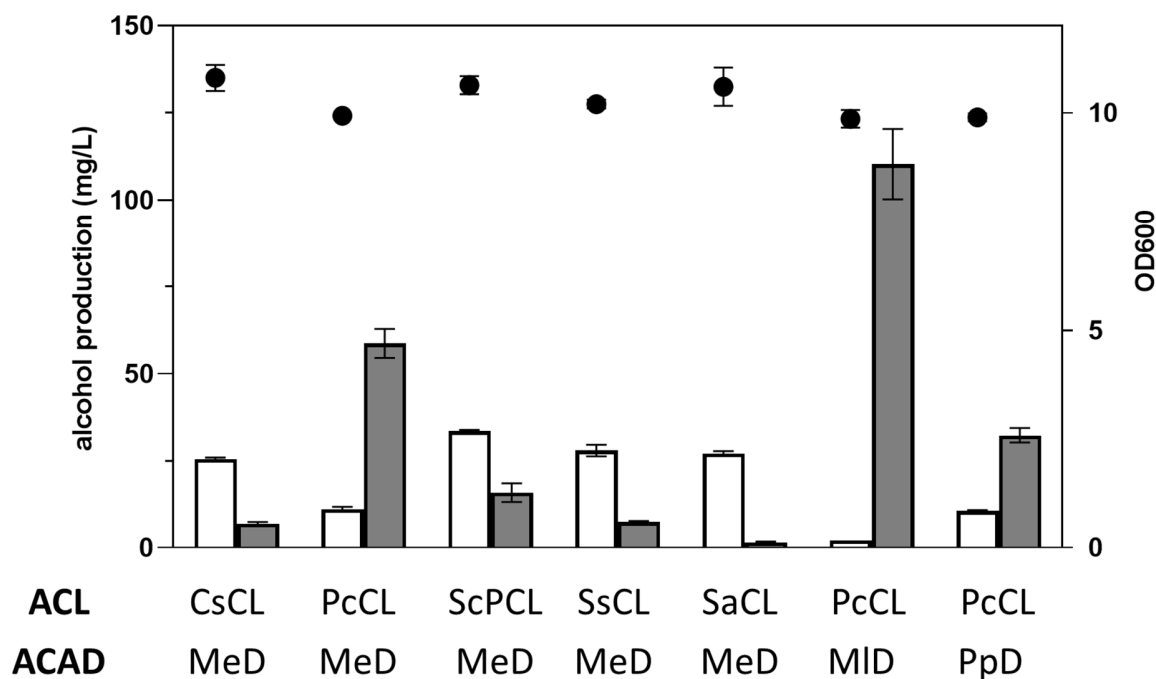|				
		chrysogenum	Hexanoate	
Acyl-CoA dehydrogenase (ACAD)				
Acyl-CoA oxidase	MlD	*Micrococcus luteus*	N/A	Goh et al. 2012
Short-chain acyl-CoA dehydrogenase	PpD	*Pseudomonas putida* KT2440	19355 Valeryl-CoA	McMahon et al. 2005
Butyryl-CoA dehydrogenase	MeD	*Megasphaera elsdenii*	10^5 Butyryl-CoA	DuPlessis et al. 1998

[*] for the substrates indicated

Fig. 3 Screening protein homologs catalyzing the first two steps of the beta-oxidation LMVA pathway. Grey bars, C6-isoprenol production; White bars, isoprenol production; Dots, endpoint optical density (OD600) of the production broth.

After optimizing the pathway genes to increase the production efficiency, we turned to the host genes that may degrade the key intermediate, 3-ketovaleryl-CoA. As mentioned before, this intermediate can be degraded by a thiolase into acetyl-CoA and propionyl-CoA. Therefore we focused on two *E. coli* chromosomal thiolase genes, *atoB* and *yqeF*, which have substrate preference for short-chain beta-ketoacyl-CoA (Dellomonaco et al., 2011; Kim et al., 2015) and were proposed to degrade 3-ketovaleryl-CoA. We conducted in-frame single-gene knockouts (Baba et al., 2006) to delete these two genes in *E. coli* BL21(DE3) in sequence, and PCR confirmed the genotypes of the knockout mutants (**Fig. S2**). The intermediate strain *E. coli* BL21(DE3) *ΔatoB* (6C01) and the final double knockout strain *E. coli* BL21(DE3) *ΔatoB ΔyqeF* (6C02) were then used for C6-isoprenol production using the plasmid combination of pJL01-MlD/pJL02-PcCL. The production of C6-isoprenol and the consumption of valeric acid were quantified. The results showed that while the knockout of *atoB* did not impact the C6-isoprenol titer, the double knockout strain, 6C02, increased the C6-isoprenol production to 390 mg/L, and the C6-isoprenol yield from valeric acid doubled to 44 mol% over the wild-type strain (**Fig. 4**). We noticed even after knocking out the thiolase genes, only around half of the valeric acid is transformed into C6-isoprenol. The loss of valeric acid may come from the evaporation and the flux into side directions in the metabolic network, such as the degradation of 3-ketovaleryl-CoA by other thiolases in *E. coli* (Campbell et al., 2003; Feigenbaum and Schulz, 1975).
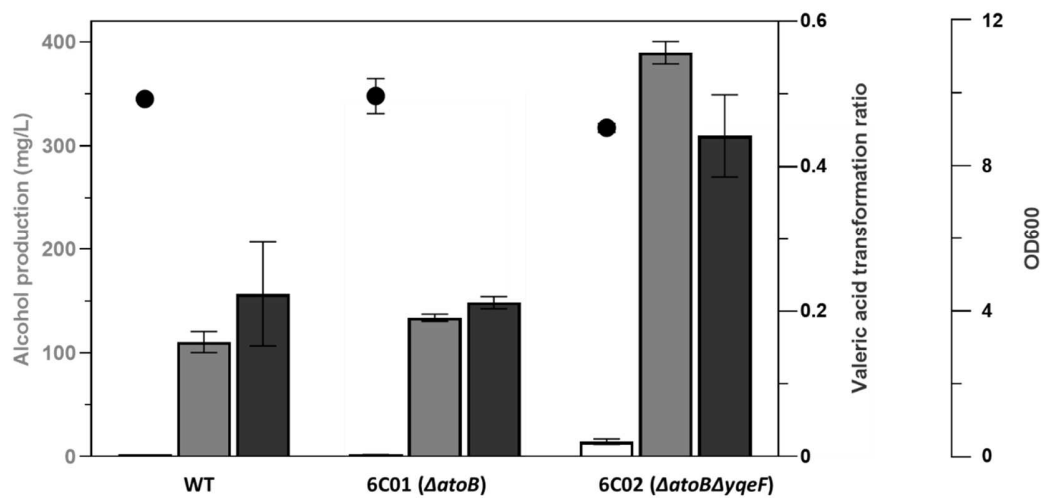
Fig. 4 Knockout short-chain acyl-CoA thiolase genes in *E. coli* BL21(DE3) to increase C6-isoprenol production and valeric acid transforming ratio. Grey bars, C6-isoprenol production; White bars, isoprenol production; Black bars, C6-isoprenol yield from valeric acid; Dots, endpoint optical density (OD600) of the cells in the production broth.

3.4 Substrate promiscuity assays of the beta-oxidation LMVA pathway

In the C6-isoprenol runs, we noticed that the levels of isoprenol were generally negatively correlated to the levels of C6-isoprenol. (**Fig.3** and **Fig.4**) For the sources of isoprenol, we reason that in addition to the *E. coli* native MEP pathway, the LMVA pathway may contribute the major portion via acetoacetyl-CoA, instead of 3-ketovaleryl-CoA, to C6-isoprenol. Through the LMVA pathway, the productions of C6-isoprenol and isoprenol use the same precursor, acetyl-CoA, resulting in the negatively correlated levels of C6-isoprenol and isoprenol. Also, the knockout of the short-chain acyl-CoA thiolase increased the supply of acetoacetyl-CoA and the production of IPP, resulting in an increased isoprenol titer of 14.4 mg/L compared to 2.1 mg/L for the wild-type strain. (**Fig. 4**)

Hence, we proposed that acetoacetyl-CoA is readily accepted by the LMVA pathway, and it's beta-oxidation precursor, butyric acid, might be a substate of the beta-oxidation LMVA pathway. To test this hypothesis, we fed 1 g/L butyric acid (C4A) instead of valeric acid to strain 6C02 containing pJL01-MlD and pJL02-PcCL. After production, GC-FID (**Fig. 5B-i,iv**) and GC-MS (**Fig. 5C-i, iv isoprenol**) confirmed the isoprenol production, quantified at 301.8 mg/L (**Fig. 5D**). This result validates butyric acid is a good substrate for the beta-oxidation LMVA pathway for IPP/isoprenol production.

The successful transformation of butyric acid into C5 alcohols by C4A suggests that the beta-oxidation LMVA pathway has substrate promiscuity. To explore the substrate spaces of this pathway, we tested other fatty acids as substrates. Without supplementing hexanoic acid, *E. coli* 6C02 with the beta-oxidation LMVA pathway also produces a small amount of C7-isoprenol, confirmed by the synthetic standard using GC-FID (**Fig. 5B-iii, iv**) and GC-MS (**Fig. 5C-iii, iv C7-isoprenol**). Supplementing hexanoic acid increased the production of C7-isoprenol significantly (**Fig. 5B-ii** and **Fig. 5D**). Therefore, the beta-oxidation LMVA pathway can activate hexanoic acid and transform it to C7-isoprenol, albeit with low efficiency. C7-isoprenol production without hexanoic acid supplementation is likely from endogenous hexanoyl-CoA in *E. coli*. We also tried fatty acid analogs with functional group substitutes, including 5-chloro-valeric acid, 4-pentenoic acid, 4-amino-butyric acid, 5,5,5-trifluorovaleric acid, and 4-bromobutyric acid. We conducted comparative GC-FID/GC-MS analysis and fragment search in GC-MS for lack of standards of the expected alcohol products. However, none of the expected substituted alcohols were detected, suggesting these fatty acid analogs are poor substrates for the beta-oxidation LMVA pathway (data not shown). The substrate promiscuity assays suggest the beta-oxidation LMVA pathway can produce IPP and C7-IPP, expanding the product chemical space of the homoterpene biosynthesis platform.
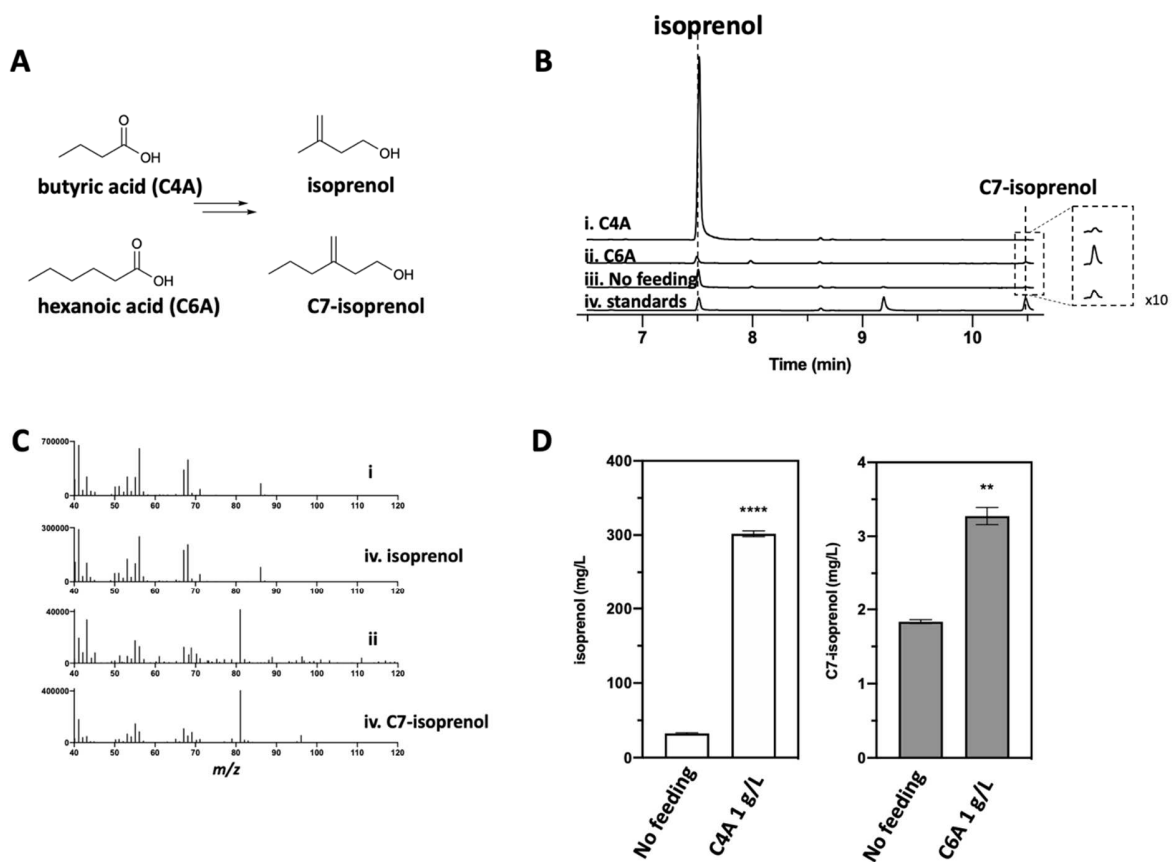
Fig. 5 Substrate promiscuity assays of the beta-oxidation LMVA pathway. (A) The beta-oxidation LMVA pathway accepts butyric acid (C4A) and hexanoic acid (C6A) into isoprenol and C7-isoprenol, respectively. (B) GC-FID traces of the production runs using *E. coli* strains coexpressing NudB and the beta-oxidation LMVA pathway with the feeding of C4A (i) and C6A (ii), along with the no feeding production run (iii) and standards (iv). (C) Mass spectrums of isoprenol and C7-isoprenol in the production broths and standards. (D) Quantification of isoprenol and C7-isoprenol in the production broths. Stars indicate significant differences between feeding and no feeding, four stars, $P \leq 0.0001$; two stars, $P \leq 0.01$.

3.5 Fuel properties of isoprenol analogs

The isopentenols, including isoprenol and prenol, are drop-in alcohol biofuels and have versatile potential fuel applications: prenol is one of the top 10 Department of Energy (DOE) Co-Optimization of Fuels & Engines (Co-Optima) gasoline blendstocks and has synergistic blending effects for research octane number (RON) (Gaspar et al., 2019; Monroe et al., 2019), and isoprenol is the precursor of 1,4-dimethylcyclooctane (DMCO), a high-performance jet fuel blendstock (Baral et al., 2021; Rosenkoetter et al., 2019). Previous studies have shown that for alcohol biofuels, molecules with longer chain lengths have a better blend stability with conventional fuels, and are less hygroscopic than their shorter chain congeners (Rajesh Kumar and Saravanan, 2016). Our pathway to C6-isoprenol and C7-isoprenol from biomass-derived fatty acids makes it possible to produce these chain-extended isoprenol analogs sustainably.

With the synthesized C6-isoprenol and C7-isoprenol, we were able to test some important fuel properties of these novel isoprenol analogs. We first estimated their water solubility based on their logP values. High water solubility contributes to the high hygroscopic nature of alcohol fuels, increasing the possibility of phase separation when blended with conventional hydrocarbon fuel. Also, for microbial biofuel production process, molecules with high logP and low water solubility will partition into the organic phase in a two-phase extractive fermentation, resulting in low product toxicity to the producing microbes (Kang et al., 2019). Increasing carbon chain length leads to decreasing polarity of alcohols, resulting in lower water solubilities. The logP of isoprenol analogs were determined using an HPLC method, with C4-C8 strain chain alcohols as references. The result revealed that the one-carbon increase of the chain length of isoprenol decreases the water solubility to 22.12 g/L from 65.36 g/L. Moreover, the addition of two carbons to isoprenol further decreases the water solubility to 6.6 g/L (**Table 2**). The decreasing trend of the water solubility was expected, and these data reflect the trend quantitatively.

The energy density of the isoprenol analogs was determined by testing their gross heats of combustion using a standard method (ASTM D4809). The energy density tests revealed C6-isoprenol has a higher heating value (HHV) of 35.524 MJ/kg, while C7-isoprenol had an HHV of 39.468 MJ/kg (**Table 2**). These numbers are similar or higher to the HHVs of isopentenols tested in the same batch. We also determined the RONs of the 10% alcohol RBOB gasoline blends using the recently published AFIDA method (Luecke and Zigler, 2021). The results indicated that C6- and C7-isoprenols have comparable RON boosting effects to isopentenols, making these two chemicals potential blendstocks for gasoline blends.

Table 2

Fuel properties of isoprenol analogs

Compound	logP	Water solubility (g/L)	HHV (MJ/kg)	10% Blended RON[*]
C7-isoprenol	2.204 ± 0.004	6.60 ± 0.076	39.468	87.5
C6-isoprenol	1.496 ± 0.004	22.12 ± 0.24	35.524	87.7
isoprenol	1.010 ± 0.005	65.36 ± 0.83	36.546	88.6
prenol	1.011 ± 0.005	65.07 ± 0.78	33.602	87.6

[*] in reformulated blendstock for oxygenate blending gasoline (RBOB); Neat RBOB RON: 86.3

4. Discussion

We previously constructed the homoterpene biosynthesis platform as a proof of concept that introduces terpene structural diversity at the precursor stage. Here we further optimized this platform towards practical application. The most significant change is the upstream pathway to the key intermediate, 3-ketovaleryl-CoA. Like the natural LMVA pathways, our previous pathway starts from propionyl-CoA, condensed by a thiolase into 3-ketovaleryl-CoA. Two points led us to consider an alternative pathway to

3-ketovaleryl-CoA. First, thermodynamic analysis indicated the condensation reaction catalyzed by thiolase is endergonic with a positive Gibbs free energy change (for acetoacetyl-CoA, ΔG = 29.6 KJ/mol) (Dellomonaco et al., 2011), suggesting the thiolase catalyzed condensation reaction is thermodynamically unfavorable. This calculation is consistent with the finding that PhaA homologs catalyze the degradation reaction better than the condensation reaction (Kim and Kim, 2014). Second, to our knowledge, almost all the reported thiolases that convert propionyl-CoA and acetyl-CoA into 3-ketovaleryl-CoA also convert two molecules of acetyl-CoA into acetoacetyl-CoA. Our later experiment using butyric acid as substrate in the beta-oxidation LMVA pathway suggests the LMVA pathway readily accepts acetoacetyl-CoA into IPP (Aldor and Keasling, 2001; Kim and Kim, 2014). Considering these two points and the relatively high concentration of acetyl-CoA in *E. coli* (Bennett et al., 2009), we reasoned that it would be difficult for the thiolase LMVA pathway to make HIPP over IPP, complicating C16, C17 and C18 terpene biosynthesis. Instead, the beta-oxidation pathway is a more specific way to produce isopentenyl pyrophosphate analogs. Although background IPP production remains, either via the native MEP pathway or from endogenous acetoacetyl-CoA transformed by the LMVA pathway, the ratio of HIPP/IPP production is significantly improved, as the molar ratio of C6-isoprenol:isoprenol is over 60 in the production run using the *E. coli* BL21(DE3) *ΔatoB* host with 1 g/L valeric acid feeding. The high production of HIPP and its dominant content in the isopentenyl pyrophosphate analog pool will benefit future homoterpene biosynthesis efforts.

Using isoprenol analogs as the final product, we successfully optimized the flux to HIPP in the homoterpene biosynthesis platform. The enzymes after the LMVA pathway leading to complex terpenes are more challenging to optimize because of their elusive enzymology and unusual substrates. Following HIPP production, an IDI is supposed to isomerize HIPP to HDMAPP. We could not detect the corresponding alcohol of HDMAPP, C6-prenol, in the *E. coli* production run with the expression of NudB and the thiolase LMVA pathway containing the IDI from *Bombyx mori*. While this IDI was confirmed *in vitro* to transform HIPP to HDMAPP specifically (Tanetoshi et al., 1985), the absence of C6-prenol in the

production run suggests that this IDI does not work well in *E. coli*, or the hydrolyzed product of HDMAPP, HDMAP, is not well accepted by the *E. coli* endogenous phosphatases. Incorporating this IDI into the optimized beta-oxidation LMVA pathway may increase HDMAPP production by increasing substrate supply. Other Lepidopteran IDIs are also candidates for enzyme screening of this step. At the same time, it is noteworthy that the regiospecificity of some Lepidopteran IDIs are low, because they transform HIPP to not only HDMAPP but also the (E)-isomer of HDMAPP and isomers with a γ-δ double bond (Sen et al., 2012). After the isomerization, ideally, one molecule of HDMAPP is supposed to condense with different molecules of HIPP into homo-GPP (1 HIPP), homo-FPP (2 HIPPs), and homo-GGPP (3 HIPPs). To our knowledge, all the characterized FPPSs that produce homo-FPP also produce FPP, with varied substrate preferences. This substrate promiscuity could explain why our previous work only produced C16 homosesquiterpenes as low, or non-existent HDMAPP levels may hamper homo-FPP analog production. The overwhelming production of HIPP to IPP in the optimized platform here may increase the HIPP incorporation to produce more C16, C17, and even C18 FPPs. For other prenyl diphosphates, reported point mutations in prenyltransferases that change the product profiles could be applied on the lepidopteran FPPS to produce homo-GPP (Zhao et al., 2017). Also, the structural basis of substrate preference for HIPP/HDMAPP derived prenyl diphosphates has been analyzed, the results of which are proposed to direct the engineering of non-lepidopteran prenyltransferases to accept HIPP/HDMAPP (Cusson et al., 2006). Finally, terpene synthases cyclized the homo prenyl pyrophosphates to terpene scaffolds. This step is the most challenging due to the lack of natural enzymes using homo prenyl pyrophosphates as the substrates. Future studies will focus on using rational design and directed evolution to alter the substrate specificity of canonical terpene synthases.

Introducing extra carbon(s) in the terpenes can significantly change their properties, exemplified by the optimized fuel properties of isoprenol biofuels. With comparable RON boosting effects and energy densities to isopentenols, C6- and C7-isoprenols have decreased water solubilities, making them better ingredients for fuel blends (Rajesh Kumar and Saravanan, 2016). In particular, C6-isoprenol derives from

valeric acid, a key intermediate in the valerate biofuel platform (Lange et al., 2010). Numerous chemical reactions/processes have been developed to transform lignocellulose to valeric acid via levulinic acid (Karanwal et al., 2020; Kon et al., 2014; Yu et al., 2019), making our pathway promising to produce C6-isoprenol as a next-generation biofuel. Besides the simple terpenes we produce here, another example is (S)-germacrene D, whose analog with two extra carbons, (S)-14,15-dimethylgermacrene D shows a reversal in insect behavioral activity (Touchet et al., 2015). Addressing those challenges in the homoterpene biosynthesis will enable the efficient production of various terpene analogs, leading to more diversified structures in the chemical portfolio for downstream applications.

Competing financial interest

J.D.K. has a financial interest in Amyris, Lygos, Demetrix, Napigen, Maple Bio, Berkeley Brewing Sciences, Ansa Biotech, Apertor Pharmaceuticals, and Zero Acre Farms. C.B.E has a financial interest in Perlumi Chemicals, Inc.

References

Ajikumar, P.K., Tyo, K., Carlsen, S., Mucha, O., Phon, T.H., Stephanopoulos, G., 2008. Terpenoids:

opportunities for biosynthesis of natural product drugs using engineered microorganisms. Mol. Pharm. 5, 167–190. doi:10.1021/mp700151b

Aldor, I., Keasling, J.D., 2001. Metabolic engineering of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) composition in recombinant *Salmonella enterica* serovar typhimurium. Biotechnol. Bioeng. 76, 108–114. doi:10.1002/bit.1150

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 2006.0008. doi:10.1038/msb4100050

Baral, N.R., Yang, M., Harvey, B.G., Simmons, B.A., Mukhopadhyay, A., Lee, T.S., Scown, C., 2021. Production cost and carbon footprint of biomass-derived dimethylcyclooctane as a high performance jet fuel blendstock. doi:10.26434/chemrxiv.14761002.v1

Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., Rabinowitz, J.D., 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. Nat. Chem. Biol. 5, 593–599. doi:10.1038/nchembio.186

Buhaescu, I., Izzedine, H., 2007. Mevalonate pathway: a review of clinical and therapeutical implications. Clin. Biochem. 40, 575–584. doi:10.1016/j.clinbiochem.2007.03.016

Campbell, J.W., Morgan-Kiss, R.M., Cronan, J.E., 2003. A new *Escherichia coli* metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway. Mol. Microbiol. 47, 793–805. doi:10.1046/j.1365-2958.2003.03341.x

Chiniquy, J., Garber, M.E., Mukhopadhyay, A., Hillson, N.J., 2020. Fluorescent amplification for next generation sequencing (FA-NGS) library preparation. BMC Genomics 21, 85. doi:10.1186/s12864-020-6481-8

Chou, H.H., Keasling, J.D., 2012. Synthetic pathway for production of five-carbon alcohols from isopentenyl diphosphate. Appl. Environ. Microbiol. 78, 7849–7855. doi:10.1128/AEM.01175-12

Cox-Georgian, D., Ramadoss, N., Dona, C., Basu, C., 2019. Therapeutic and medicinal uses of terpenes, in: Joshee, N., Dhekney, S.A., Parajuli, P. (Eds.), Medicinal Plants: From Farm to Pharmacy.

Springer International Publishing, Cham, pp. 333–359. doi:10.1007/978-3-030-31269-5_15

Cusson, M., Béliveau, C., Sen, S.E., Vandermoten, S., Rutledge, R.G., Stewart, D., Francis, F., Haubruge, E., Rehse, P., Huggins, D.J., Dowling, A.P.G., Grant, G.H., 2006. Characterization and tissue-specific expression of two lepidopteran farnesyl diphosphate synthase homologs: implications for the biosynthesis of ethyl-substituted juvenile hormones. Proteins 65, 742–758. doi:10.1002/prot.21057

Dellomonaco, C., Clomburg, J.M., Miller, E.N., Gonzalez, R., 2011. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. Nature 476, 355–359. doi:10.1038/nature10333

Dunlop, M.J., 2011. Engineering microbes for tolerance to next-generation biofuels. Biotechnol. Biofuels 4, 32. doi:10.1186/1754-6834-4-32

DuPlessis, E.R., Pellett, J., Stankovich, M.T., Thorpe, C., 1998. Oxidase activity of the acyl-CoA dehydrogenases. Biochemistry 37, 10469–10477. doi:10.1021/bi980767s

Eiben, C.B., de Rond, T., Bloszies, C., Gin, J., Chiniquy, J., Baidoo, E.E.K., Petzold, C.J., Hillson, N.J., Fiehn, O., Keasling, J.D., 2019. Mevalonate pathway promiscuity enables noncanonical terpene production. ACS Synth. Biol. 8, 2238–2247. doi:10.1021/acssynbio.9b00230

Feigenbaum, J., Schulz, H., 1975. Thiolases of Escherichia coli: purification and chain length specificities. J. Bacteriol. 122, 407–411. doi:10.1128/jb.122.2.407-411.1975

Gaspar, D.J., West, B.H., Ruddy, D., Wilke, T.J., Polikarpov, E., Alleman, T.L., George, A., Monroe, E., Davis, R.W., Vardon, D., Sutton, A.D., Moore, C.M., Benavides, P.T., Dunn, J., Biddy, M.J., Jones, S.B., Kass, M.D., Pihl, J.A., Pihl, J.A., Debusk, M.M., Pitz, W.J., 2019. Top ten blendstocks derived from biomass for turbocharged spark ignition engines: bio-blendstocks with potential for highest engine efficiency. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States). doi:10.2172/1567705

Goh, E.-B., Baidoo, E.E.K., Keasling, J.D., Beller, H.R., 2012. Engineering of bacterial methyl ketone synthesis for biofuels. Appl. Environ. Microbiol. 78, 70–80. doi:10.1128/AEM.06785-11

Go, M.K., Chow, J.Y., Cheung, V.W.N., Lim, Y.P., Yew, W.S., 2012. Establishing a toolkit for

precursor-directed polyketide biosynthesis: exploring substrate promiscuities of acid-CoA ligases. Biochemistry 51, 4568–4579. doi:10.1021/bi300425j

Hillson, N.J., Rosengarten, R.D., Keasling, J.D., 2012. j5 DNA assembly design automation software. ACS Synth. Biol. 1, 14–21. doi:10.1021/sb2000116

Huang, Z.-R., Lin, Y.-K., Fang, J.-Y., 2009. Biological and pharmacological activities of squalene and related compounds: potential uses in cosmetic dermatology. Molecules 14, 540–554. doi:10.3390/molecules14010540

Kang, A., Mendez-Perez, D., Goh, E.-B., Baidoo, E.E.K., Benites, V.T., Beller, H.R., Keasling, J.D., Adams, P.D., Mukhopadhyay, A., Lee, T.S., 2019. Optimization of the IPP-bypass mevalonate pathway and fed-batch fermentation for the production of isoprenol in *Escherichia coli*. Metab. Eng. 56, 85–96. doi:10.1016/j.ymben.2019.09.003

Karanwal, N., Verma, D., Butolia, P., Kim, S.M., Kim, J., 2020. One-pot direct conversion of levulinic acid into high-yield valeric acid over a highly stable bimetallic Nb-Cu/Zr-doped porous silica catalyst. Green Chem. 22, 766–787. doi:10.1039/C9GC03516H

Kim, E.-J., Kim, K.-J., 2014. Crystal structure and biochemical characterization of PhaA from *Ralstonia eutropha*, a polyhydroxyalkanoate-producing bacterium. Biochem. Biophys. Res. Commun. 452, 124–129. doi:10.1016/j.bbrc.2014.08.074

Kim, S., Clomburg, J.M., Gonzalez, R., 2015. Synthesis of medium-chain length (C6-C10) fuels and chemicals via β-oxidation reversal in *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 42, 465–475. doi:10.1007/s10295-015-1589-6

Koetsier, M.J., Jekel, P.A., Wijma, H.J., Bovenberg, R.A.L., Janssen, D.B., 2011. Aminoacyl-coenzyme A synthesis catalyzed by a CoA ligase from *Penicillium chrysogenum*. FEBS Lett. 585, 893–898. doi:10.1016/j.febslet.2011.02.018

Kon, K., Onodera, W., Shimizu, K., 2014. Selective hydrogenation of levulinic acid to valeric acid and valeric biofuels by a Pt/HMFI catalyst. Catal. Sci. Technol. 4, 3227–3234. doi:10.1039/C4CY00504J

Lange, J.-P., Price, R., Ayoub, P.M., Louis, J., Petrus, L., Clarke, L., Gosselink, H., 2010. Valeric

biofuels: a platform of cellulosic transportation fuels. Angew. Chem. Int. Ed 49, 4479–4483. doi:10.1002/anie.201000655

Luecke, J., Zigler, B.T., 2021. Rapid prediction of fuel research octane number and octane sensitivity using the AFIDA constant-volume combustion chamber. Fuel 301, 120969. doi:10.1016/j.fuel.2021.120969

Luo, X., Reiter, M.A., d'Espaux, L., Wong, J., Denby, C.M., Lechner, A., Zhang, Y., Grzybowski, A.T., Harth, S., Lin, W., Lee, H., Yu, C., Shin, J., Deng, K., Benites, V.T., Wang, G., Baidoo, E.E.K., Chen, Y., Dev, I., Petzold, C.J., Keasling, J.D., 2019. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. Nature 567, 123–126. doi:10.1038/s41586-019-0978-9

McMahon, B., Gallagher, M.E., Mayhew, S.G., 2005. The protein coded by the PP2216 gene of *Pseudomonas putida* KT2440 is an acyl-CoA dehydrogenase that oxidises only short-chain aliphatic substrates. FEMS Microbiol. Lett. 250, 121–127. doi:10.1016/j.femsle.2005.06.049

Men, X., Wang, F., Chen, G.-Q., Zhang, H.-B., Xian, M., 2018. Biosynthesis of natural rubber: current state and perspectives. Int. J. Mol. Sci. 20. doi:10.3390/ijms20010050

Miyazawa, T., Takahashi, S., Kawata, A., Panthee, S., Hayashi, T., Shimizu, T., Nogawa, T., Osada, H., 2015. Identification of middle chain fatty acyl-CoA ligase responsible for the biosynthesis of 2-alkylmalonyl-CoAs for polyketide extender unit. J. Biol. Chem. 290, 26994–27011. doi:10.1074/jbc.M115.677195

Monroe, E., Gladden, J., Albrecht, K.O., Bays, J.T., McCormick, R., Davis, R.W., George, A., 2019. Discovery of novel octane hyperboosting phenomenon in prenol biofuel/gasoline blends. Fuel 239, 1143–1148. doi:10.1016/j.fuel.2018.11.046

Oldfield, E., Lin, F.-Y., 2012. Terpene biosynthesis: modularity rules. Angew. Chem. Int. Ed 51, 1124–1137. doi:10.1002/anie.201103110

Pfeifer, B.A., Admiraal, S.J., Gramajo, H., Cane, D.E., Khosla, C., 2001. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. Science 291, 1790–1792. doi:10.1126/science.1058092

Rajesh Kumar, B., Saravanan, S., 2016. Use of higher alcohol biofuels in diesel engines: A review. Renew. Sustain. Energy Rev 60, 84–115. doi:10.1016/j.rser.2016.01.085

Rivas, F., Parra, A., Martinez, A., Garcia-Granados, A., 2013. Enzymatic glycosylation of terpenoids. Phytochem. Rev. 12, 327–339. doi:10.1007/s11101-013-9301-9

Rohmer, M., 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. Nat. Prod. Rep. 16, 565–574. doi:10.1039/a709175c

Rosenkoetter, K.E., Kennedy, C.R., Chirik, P.J., Harvey, B.G., 2019. [4 + 4]-cycloaddition of isoprene for the production of high-performance bio-based jet fuel. Green Chem. 21, 5616–5623. doi:10.1039/C9GC02404B

Schooley, D.A., Judy, K.J., Bergot, B.J., Hall, M.S., Siddall, J.B., 1973. Biosynthesis of the Juvenile Hormones of *Manduca sexta*: Labeling Pattern from Mevalonate, Propionate, and Acetate. Proc Natl Acad Sci USA 70, 2921–2925. doi:10.1073/pnas.70.10.2921

Sen, S.E., Tomasello, A., Grasso, M., Denton, R., Macor, J., Béliveau, C., Cusson, M., Crowell, D.N., 2012. Cloning, expression and characterization of lepidopteran isopentenyl diphosphate isomerase. Insect Biochem. Mol. Biol. 42, 739–750. doi:10.1016/j.ibmb.2012.07.001

Stout, J.M., Boubakir, Z., Ambrose, S.J., Purves, R.W., Page, J.E., 2012. The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. Plant J. 71, 353–365. doi:10.1111/j.1365-313X.2012.04949.x

Tanetoshi, K., Minoru, M., Kyozo, O., 1985. Isoprenoid Enzyme Systems of Silkworm II. Formation of the Juvenile Hormone Skeletons by Farnesyl Pyrophosphate Synthetase II. The Journal of Biochemistry.

Thompson, M.G., Sedaghatian, N., Barajas, J.F., Wehrs, M., Bailey, C.B., Kaplan, N., Hillson, N.J., Mukhopadhyay, A., Keasling, J.D., 2018. Isolation and characterization of novel mutations in the pSC101 origin that increase copy number. Sci. Rep. 8, 1590. doi:10.1038/s41598-018-20016-w

Touchet, S., Chamberlain, K., Woodcock, C.M., Miller, D.J., Birkett, M.A., Pickett, J.A., Allemann, R.K., 2015. Novel olfactory ligands via terpene synthases. Chem. Commun 51, 7550–7553.

doi:10.1039/c5cc01814e

Tu, Q., Yin, J., Fu, J., Herrmann, J., Li, Y., Yin, Y., Stewart, A.F., Müller, R., Zhang, Y., 2016. Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. Sci. Rep. 6, 24648. doi:10.1038/srep24648

Tu, Y., 2011. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. Nat. Med. 17, 1217–1220. doi:10.1038/nm.2471

Wenig, P., Odermatt, J., 2010. OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data. BMC Bioinformatics 11, 405. doi:10.1186/1471-2105-11-405

Yu, Z., Lu, X., Xiong, J., Ji, N., 2019. Transformation of levulinic acid to valeric biofuels: A review on heterogeneous bifunctional catalytic systems. ChemSusChem 12, 3915–3930. doi:10.1002/cssc.201901522

Zada, B., Wang, C., Park, J.-B., Jeong, S.-H., Park, J.-E., Singh, H.B., Kim, S.-W., 2018. Metabolic engineering of *Escherichia coli* for production of mixed isoprenoid alcohols and their derivatives. Biotechnol. Biofuels 11, 210. doi:10.1186/s13068-018-1210-0

Zhang, J., Barajas, J.F., Burdu, M., Wang, G., Baidoo, E.E., Keasling, J.D., 2017. Application of an acyl-CoA ligase from *Streptomyces aizunensis* for lactam biosynthesis. ACS Synth. Biol. 6, 884–890. doi:10.1021/acssynbio.6b00372

Zhao, J., Li, C., Zhang, Y., Shen, Y., Hou, J., Bao, X., 2017. Dynamic control of ERG20 expression combined with minimized endogenous downstream metabolism contributes to the improvement of geraniol production in *Saccharomyces cerevisiae*. Microb. Cell Fact. 16, 17. doi:10.1186/s12934-017-0641-9

Zheng, X., Li, P., Lu, X., 2019. Research advances in cytochrome P450-catalysed pharmaceutical terpenoid biosynthesis in plants. J. Exp. Bot. 70, 4619–4630. doi:10.1093/jxb/erz203