

<https://doi.org/10.1038/s41540-024-00440-7>

Systems biology approach for enhancing limonene yield by re-engineering *Escherichia coli*



Jasmeet Kaur Khanijou¹, Yan Ting Hee², Clement P. M. Scipion³, Xixian Chen¹ & Kumar Selvarajoo^{2,4,5}✉

Engineered microorganisms have emerged as viable alternatives for limonene production. However, issues such as low enzyme abundance or activities, and regulatory feedback/forward inhibition may reduce yields. To understand the underlying metabolism, we adopted a systems biology approach for an engineered limonene-producing *Escherichia coli* strain K-12 MG1655. Firstly, we generated time-series metabolomics data and, secondly, developed a dynamic model based on enzyme dynamics to track the native metabolic networks and the engineered mevalonate pathway. After several iterations of model fitting with experimental profiles, which also included ¹³C-tracer studies, we performed in silico knockouts (KOs) of all enzymes to identify bottleneck(s) for optimal limonene yields. The simulations indicated that ALDH/ADH (aldehyde dehydrogenase/alcohol dehydrogenase) and LDH (lactate dehydrogenase) suppression, and HK (hexokinase) enhancement would increase limonene yields. Experimental confirmation was achieved, where ALDH-ADH and LDH KOs, and HK overexpression improved limonene yield by 8- to 11-fold. Our systems biology approach can guide microbial strain re-engineering for optimal target production.

Isoprenoids are a class of plant derived natural products which are used in the pharmaceutical and food industries¹. Limonene is a monoterpene used as a flavour or fragrance based on its citrus aroma^{2,3} while its hydrogenated form is used as a biofuel⁴. Limonene exists in nature as two enantiomers, (D)-limonene and (L)-limonene, due to a chiral centre in its structure. Limonene is also a vital precursor to several compounds such as menthol, carveol, carvone, and perillyl alcohol, which are used in food and beverage, pharmaceuticals, and cosmetics⁵. Currently, limonene is obtained through the extraction of plant biomass, which is rather inefficient due to variations in seasons and conditions of agricultural land⁶. With the recent progress in metabolic engineering and synthetic biology, microbes, such as *Escherichia coli*, have emerged as bio-factories providing alternative routes to produce target terpenoids such as limonene^{3,7–10}.

Escherichia coli have three glycolytic pathways, namely the Embden–Meyerhoff–Parnass (EMP), Entner–Doudoroff (ED), and pentose phosphate (PP) pathways¹¹ (Fig. 1). Glucose metabolism depends on the EMP and PP pathways while the ED pathway tends to be inactive except during growth with gluconate^{12,13}. For the production of isoprenoids in *E.*

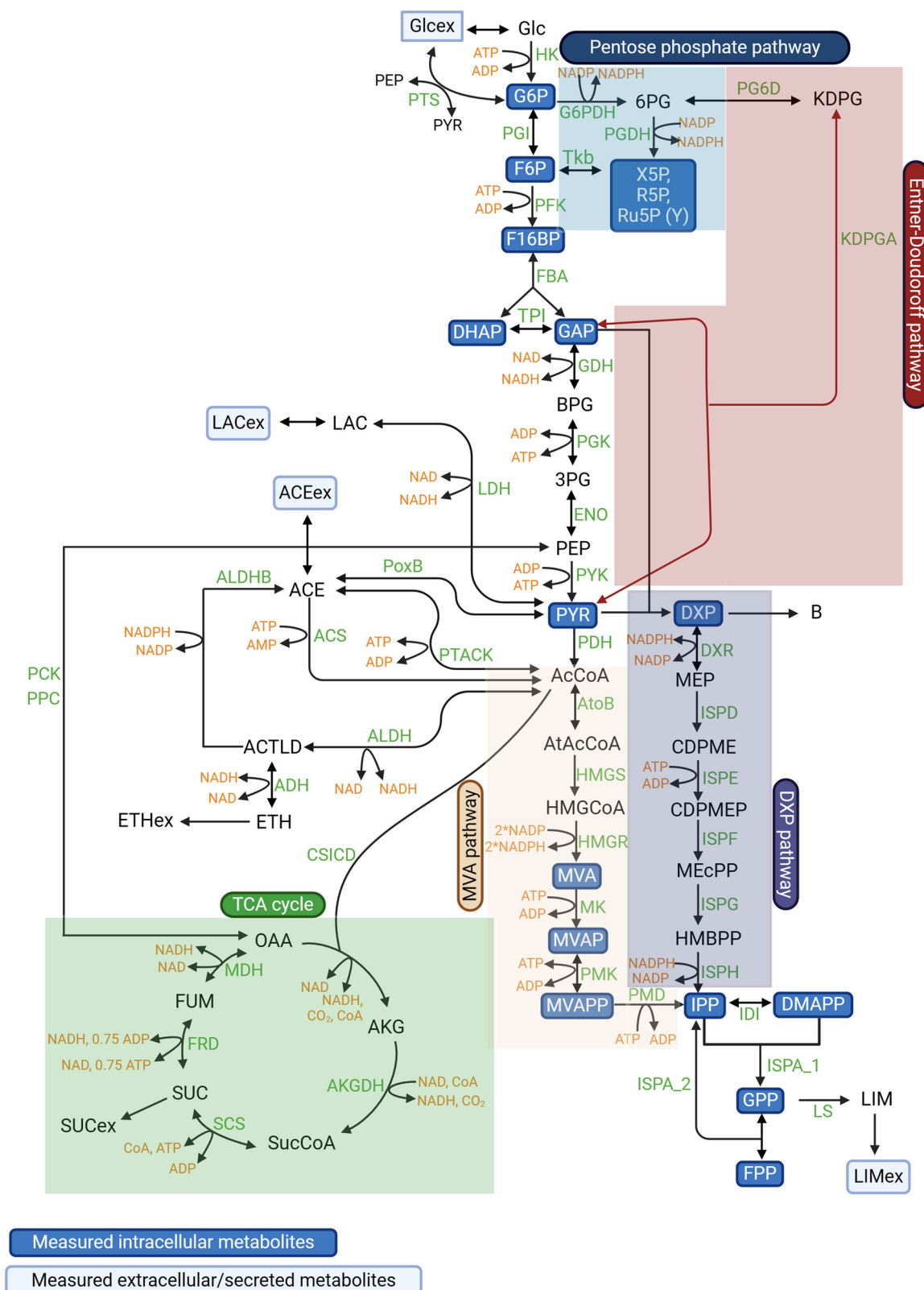
coli, there are two main pathways: the native deoxyxylulose 5-phosphate (DXP) and the heterologous mevalonate (MVA) pathways (Fig. 1). The MVA and DXP pathways both produce the isoprene building blocks isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), which are isomers used to biosynthesize isoprenoids¹⁴. To attain high isoprenoid yields in *E. coli*, the availability of co-factors, energy demands, and carbon flux need to be balanced in the DXP and MVA pathways^{15–17}.

Limonene biosynthesis has been generally low for viable bioprocesses¹⁸, and this can be attributed to several issues, such as nominal metabolic fluxes flowing towards limonene production, enzyme inefficiency in metabolic biosynthesis pathways, and limonene cytotoxicity distressing the microbial chassis¹⁹. Although the native DXP pathway results in an enhanced theoretical yield, metabolic engineering of the DXP pathway in *E. coli* for improved terpenes yields has been challenging. This is probably due to pathway regulations and minimal concentrations of isoprenoid precursors²⁰, especially geranyl pyrophosphate (GPP)⁷. Although gene modulation of the DXP pathway has improved isoprenoid production²¹, engineering a heterologous MVA pathway for expression alongside the

¹Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 31 Biopolis Way, Nanos, Singapore, 138669, Singapore. ²Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis St, Matrix, Singapore, 138671, Singapore.

³CNRS@CREATE, 1 Create Way, #08-01 Create Tower, Singapore, Singapore. ⁴Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore (NUS), Singapore, Singapore. ⁵School of Biological Sciences, Nanyang Technological University (NTU), Singapore, Singapore.

✉ e-mail: kumar_selvarajoo@bii.a-star.edu.sg



native DXP pathway in *E. coli* has improved yields of target compounds through increased precursor supply^{3,9,20,22}. In attempts to improve limonene yield in bacterial bio-factories, transcriptional tuning²³, promoter optimisation²⁴, and incorporating alternative enzymes³ have been explored. However, such strategies to improve limonene yield do not assess the involvement of competing pathways²⁵ that affect the carbon flux entering

the DXP and MVA pathways, thereby affecting limonene yield. Understanding the presence of metabolic bottlenecks and the loss of carbon flux²⁶ are also vital in enhancing limonene yields.

Metabolic networks are dynamically regulated and complex with highly interwoven networks. Due to this intricacy, metabolic engineering strategies are usually non-intuitive, where metabolic models such as

Fig. 1 | Final metabolic network topology utilised in model development of wild-type *E. coli* engineered with the mevalonate (MVA) pathway to overproduce limonene. There are several different metabolic pathways involved in limonene biosynthesis, including the glycolytic pathways, tricarboxylic acid cycle (TCA), the native deoxyxylulose 5-phosphate (DXP), and MVA pathways. Co-factor (in orange) consumption is represented by curved arrows. Intermediates: Glcex extracellular glucose, Glc glucose, G6P glucose-6-phosphate, 6PG 6-phosphogluconate, KDPG 2-keto-deoxy-6-phosphogluconate, X5P xylulose-5-phosphate, Ru5P ribulose-5-phosphate, R5P ribose-5-phosphate, F6P fructose-6-phosphate, F16BP fructose-1,6-bisphosphate, GAP glyceraldehyde-3-phosphate, DHAP dihydroxyacetone phosphate, BPG 1,3-bisphosphoglycerate, 3PG 3-phosphoglycerate, PEP phosphoenolpyruvate, PYR pyruvate, DXP 1-deoxy-D-xylulose-5-phosphate, B flux to vitamin B6 pathway, MEP 2-C-methylerythritol-4-phosphate, CDPME 4-diphosphocytidyl-2-C-methylerythritol, CDPMEP 4-diphosphocytidyl-2-C-methylerythritol-2-phosphate, MEcPP 2-C-Methylerythritol-2,4-cyclodiphosphate, HMBPP hydroxymethylbutenyl 4-diphosphate, IPP isopentenyl diphosphate, DMAPP dimethylallyl diphosphate, GPP geranyl diphosphate, FPP farnesyl diphosphate, LIM limonene, LIMex extracellular limonene, AcCoA acetyl coenzyme A, AtAcCoA acetoacetyl-CoA, HMGCoA hydroxymethylglutaryl-CoA, MVA mevalonate, MVAP 5-phosphomevalonate, MVAPP 5-diphosphomevalonate, ACE acetic acid, ACEex extracellular acetic acid, ACTLD acetaldehyde, ETH ethanol, ETHex extracellular ethanol, LAC lactic acid, LACex extracellular lactic acid, AKG α -ketoglutarate, SucCoA, succinyl-CoA, SUC succinate, SUCex extracellular

succinate, FUM fumarate, OAA oxaloacetate. Enzymes (in green): PTS phosphotransferase system, HK hexokinase, G6PDH lumped reactions of glucose-6-phosphate dehydrogenase and 6-phosphogluconolactonase, PGDH 6-phosphogluconate dehydrogenase, KDPGA 2-keto-deoxy-6-phosphogluconate aldolase, Tkb transketolase, PGI phosphoglucose isomerase, PFK phosphofructokinase, FBA fructose-1,6-bisphosphate aldolase, GDH glutamate dehydrogenase, PGK phosphoglycerate kinase, ENO enolase, PYK pyruvate kinase, DXS DXP synthase, DXR DXP reductase, ISPD CDPME synthase, ISPE CDPME kinase, ISPF MEcPP synthase, ISPG HMBPP synthase, ISPH HMBPP reductase, IDI isopentenyl diphosphate isomerase, ISPA farnesyl diphosphate synthase, LS limonene synthase, PDH pyruvate dehydrogenase, AtoB acetyl-CoA acetyltransferase, HMGS HMGCoA synthase, HMGR HMGCoA reductase, MK mevalonate kinase, PMK phosphomevalonate kinase, PMD diphosphate mevalonate decarboxylase, LDH lactate dehydrogenase, PoxB pyruvate oxidase, PCK phosphoenolpyruvate carboxykinase, PPC phosphoenolpyruvate carboxylase, ACS acetyl-CoA synthetase, PTACK lumped reactions of phosphate acetyltransferase and acetate kinase, ALDHB aldehyde dehydrogenase B, ALDH aldehyde dehydrogenase, ADH alcohol dehydrogenase, CSICD lumped enzymatic reactions of citrate synthase, aconitate hydratase A, aconitate hydratase B and isocitrate dehydrogenase, AKGDH α -ketoglutarate dehydrogenase, SCS succinyl-CoA synthetase, FRD fumarate reductase, MDH malate dehydrogenase. The illustration was created using Biorender.com.

constraint-based models and dynamic models can aid in addressing certain engineering questions²⁷. Constraint-based models, which rely on steady-state assumption, are founded on stoichiometry, and do not require mechanistic information²⁸. Thus, unlike dynamic or kinetic models, constraint-based models do not provide time-dependent information on cellular metabolism^{29,30}. Dynamic models can reveal the relationship between metabolic fluxes, enzyme expression, metabolite concentrations, and regulation via mechanistic associations, thereby granting dynamic models of metabolism vital for comprehending, predicting, and optimising metabolic behaviours^{31,32}. Dynamic models are created using linear and non-linear differential equations which are based on enzyme dynamics for each biochemical reaction within the metabolic network. These dynamic models have been utilised to understand aspects of systems biology such as metabolism and metabolic engineering^{33,34}, protein signalling^{35–38}, and gene regulatory networks^{39–41}. Dynamic models created using time-series experimental data can, thus, provide better mechanistic understanding of biological networks.

In a recent study, modelling of enzyme dynamics in photosynthetic cyanobacteria with the DXP pathway had revealed limonene synthase, which catalyses the last reaction in limonene biosynthesis, as a key limiting step in limonene bioproduction⁴². The authors achieved higher limonene production in cyanobacteria by re-engineering the limonene synthase, compared to a previous study where multiple genes were engineered laboriously based on an experimental approach alone⁴³. From this, we see that computational modelling together with synthetic biology-based engineering can be a good strategy in identifying important flux-controlling points in metabolic pathways. A systems biology approach which involves an integration of experimental research and computational modelling can also be used to understand intricate biological systems, such as engineered bacterial bio-factories. Systems biology needs to be predictive, where hypotheses are refuted or confirmed through targeted system perturbation⁴⁴.

The use of computational models, which is one of the defining features of system biology, is vital in describing the inherent intricacy of biological systems⁴⁵. Constraint-based models, such as genome-scale models, can generate pathway-level predictions which allow the identification and modification of pathway fluxes to improve the production of target products from engineered microbes. For instance, strain designs for lactic acid overproduction have been predicted using a genome-scale model of *E. coli*⁴⁶. Furthermore, algorithms such as OptKnock⁴⁷, OptStrain⁴⁸, OptORF⁴⁹, OptFlux⁵⁰, OptForce⁵¹, k-OptForce⁵², and DySScO⁵³ have been developed to assist the process of microbial strain design and expand the predictive means

of constraint-based models. Although these approaches predict the impact of genetic modifications on the microbial production of target products, the translation of these computationally predicted improved microbial strain designs and optimal flux states to the in vivo system has been rather challenging⁵⁴. To enhance the predictive capability and accuracy of constraint-based models, cellular details such as signalling pathways⁵⁵, thermodynamics⁵⁶, and transcriptional regulation⁵⁷ have been recently explored.

Dynamic or kinetic models, on the other hand, utilise more mechanistic details in a time-series manner which could potentially prove to be more useful in designing microbial strains with improved titres. Furthermore, dynamic models based on the model organism *E. coli* can be easily created as many of the metabolic reactions involving substrates, products, enzymes, and co-factors are well-known and characterised. Nevertheless, the challenge for dynamic models is in the determination of appropriate rate laws and the corresponding parameter values reliably, as many of the metabolites measured are either derived from steady-state conditions or produced extracellularly^{58,59}.

In this paper, a wild-type *E. coli* strain was first engineered to produce L-limonene through the expression of the MVA pathway and a limonene synthase (EcoCTs03)³. In parallel, a comprehensive dynamic model, based on enzyme dynamics and differential equations, was developed for the limonene biosynthesis network; by considering the numerous metabolic pathways such as the native glycolytic, TCA, mixed fermentation, and DXP pathways as well as the engineered MVA pathway. Notably, to overcome some of the challenges of dynamic models mentioned earlier, here both intracellular and extracellular time-series quantitative metabolomics data from cell cultures of EcoCTs03 were used to calibrate the dynamic model. [1-¹³C]-glucose tracer was also used to determine the extent of the ED pathway contributing to the upstream glycolytic pathway⁶⁰. The metabolic network topology (Fig. 1) and dynamic model were constructed using Complex Pathway Simulator (COPASI)⁶¹, where the collected time-series data was fed into the model for parameter estimation and simulations. The calibrated model was next used to identify intracellular targets for optimising the production of limonene for the strain. Finally, the model predictions were compared with experimental yields, by generating new strains suggested by the model.

Results

Generation of time-series metabolomics data

The engineered wild-type *E. coli*, EcoCTs03, was cultured in shake flasks and sacrificed in duplicates over various time points post IPTG induction.

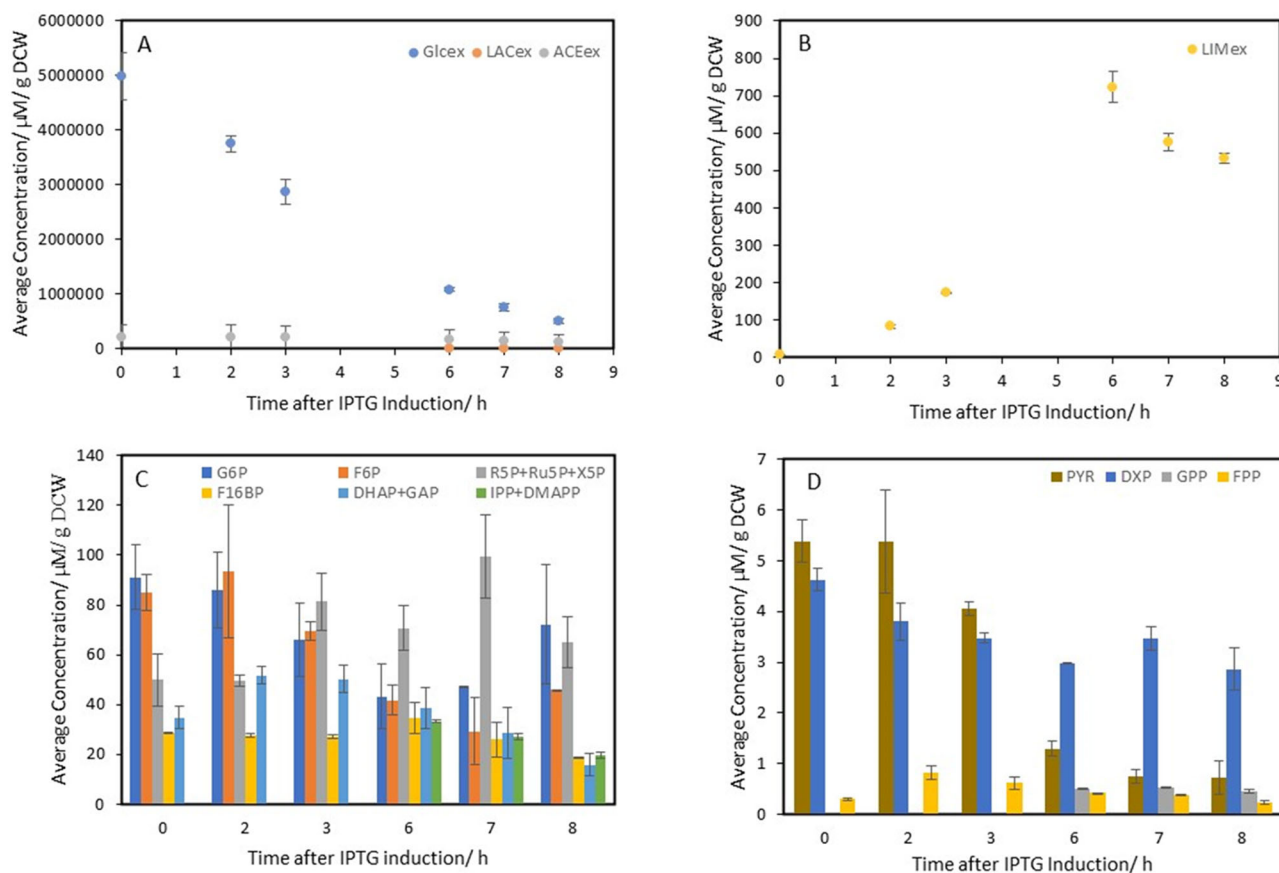


Fig. 2 | Extracellular and intracellular time-series metabolomics data obtained from cell cultures of wild-type *E. coli* engineered to overproduce limonene (EcoTs03) used for model development. **A** Reflects the concentrations of extracellular metabolites glucose (Glcex), lactic acid (LACex), and acetic acid (ACEex). **B** Displays the time-series concentrations of secreted limonene (LIMex). **C** Shows the concentrations of intracellular metabolites glucose-6-phosphate (G6P), fructose-6-phosphate (F6P), fructose-1,6-biphosphate (F16BP), DHAP + GAP (pool of dihydroxyacetone phosphate and glyceraldehyde-3-phosphate), R5P +

Ru5P + X5P (pool of ribose-5-phosphate, ribulose-5-phosphate, and xylulose-5-phosphate), and IPP + DMAPP (pool of isopentenyl pyrophosphate and dimethylallyl pyrophosphate) which are of higher concentrations. **D** Illustrates the intracellular time-series concentrations of pyruvate (PYR), deoxyxylulose 5-phosphate (DXP), geranyl diphosphate (GPP), and farnesyl diphosphate (FPP), which are of a lower concentration range. Average values are shown for biological replicates (± 1 SD).

Intracellular metabolites of the cell cultures were analysed on the LC-TOF after extraction following fast filtration and quenching using liquid nitrogen. Extracellular metabolites were analysed on the HPLC after centrifugation of the cell cultures while the dodecane layer of each cell culture was collected and subjected to GC-MS analysis for limonene quantitation. Figure 2 reflects the time-series metabolomics data collected from cultures of EcoTs03.

The extracellular metabolite glucose (Glcex) in the culture media decreased with time while the extracellular acetic acid (ACEex) concentration remained relatively constant (Fig. 2A). Extracellular lactic acid (LACex) was not observed at the initial time points of 0–3 h post IPTG induction, however it was produced and detected from 6 to 8 h post IPTG induction (Fig. 2A). Limonene was secreted extracellularly (LIMex) and trapped in the dodecane layer, where its concentration was found to generally increase with time (Fig. 2B). The changes in and around the 6th hour of IPTG induction could also be linked to the *E. coli* doubling time, as shown in Supplementary Fig. 1.

Figure 2C mainly shows the concentrations of the intracellular metabolites through the glycolytic pathways until pyruvate at the various time points. Intracellular metabolites G6P (glucose-6-phosphate) and F6P (fructose-6-phosphate) concentrations were relatively constant from 0 to 3 h post IPTG induction and dropped at 6 h post IPTG induction and were relatively constant until 8 h post IPTG induction (Fig. 2C). Intracellular R5P + Ru5P + X5P (pool of ribose-5-phosphate, ribulose-5-phosphate,

and xylulose-5-phosphate) concentrations were found to be higher at the later time points of 3 h to 8 h compared to 0 h to 2 h post IPTG induction (Fig. 2C). For the intracellular pool of dihydroxyacetone phosphate and glyceraldehyde-3-phosphate (DHAP + GAP), the concentration increased from 0 to 2 h post IPTG induction and then dropped continuously until 8 h post IPTG induction while intracellular F16BP (fructose-1,6-biphosphate) remained relatively constant through the time points (Fig. 2C). Intracellular PYR (pyruvate) remained relatively constant at 0 h and 2 h and decreased from 3 h to 8 h post IPTG induction (Fig. 2D).

Intracellular metabolites MVA (mevalonate) and MVAP (5-phosphomevalonate) were not observed in the bacterial cells during the harvesting time points. Intracellular MVAPP (5-pyrophosphomevalonate), on the other hand, was only detected at one time point of 8 h post IPTG induction at a low concentration of 0.05 ± 0.02 μM/g DCW. Figure 2D illustrates the time-series concentrations of the downstream intracellular metabolites through the MVA pathway and the intracellular metabolite DXP (1-deoxy-D-xylulose-5-phosphate). Intracellular FPP (farnesyl diphosphate) and DXP remained relatively constant through the time points (Fig. 2D). Intracellular GPP (geranyl diphosphate) was not detected at earlier time points of 0 h to 3 h, while it was detected and quantified from 6 h to 8 h post IPTG induction (Fig. 2D). Intracellular IPP + DMAPP (pool of isopentenyl pyrophosphate and dimethylallyl pyrophosphate) was found at high concentrations at later time points of 6 h to 8 h post IPTG induction and was not observed at the earlier time points (Fig. 2C).

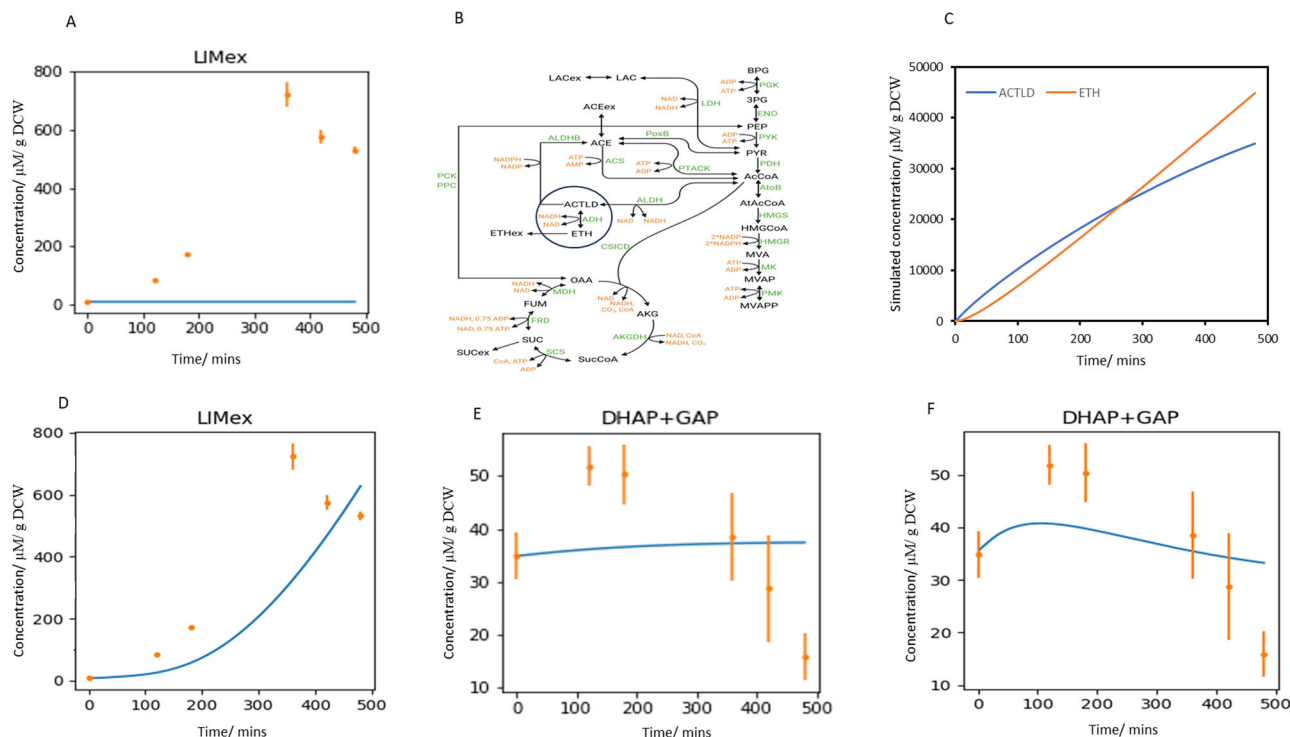


Fig. 3 | Simulations from Models A, B, and C during model iterations showing improvements in the fitting of the simulations of extracellular limonene (LIMex) and the pooled metabolites DHAP + GAP (pool of dihydroxyacetone phosphate and glyceraldehyde-3-phosphate) with experimental results. A Illustrates the poor fitting of LIMex simulation from Model A with the experimental time-series data. **B** Displays the metabolic network topology where acetaldehyde (ACTLD) and

ethanol (ETH) reside, while **C** shows the simulated ACTLD and ETH concentrations with time from Model A. **D** Shows an improved LIMex simulation from Model B while **E** reflects the poor fitting of DHAP + GAP simulation from Model B. **F** Displays the improved simulation of DHAP + GAP from Model C with the Entner–Doudoroff (ED) pathway included in the metabolic network. Average values are shown for biological replicates (± 1 SD) in (A, D–F).

The extracellular metabolite Glcex decreased with time whilst ACEex and LACex were detected and quantified from the cell culture media. LIMex was secreted in the dodecane layer and generally increased with time. The intracellular metabolites involved in the upper glycolytic pathways (G6P, F6P, F16BP, DHAP + GAP, and R5P + Ru5P + X5P) were found to be in higher concentrations compared to intracellular PYR since it marks the endpoint of glycolysis. Furthermore, the intracellular metabolites involved in the upper glycolytic pathways (G6P, F6P, F16BP, DHAP + GAP and R5P + Ru5P + X5P) also had higher concentrations compared to the intracellular metabolites found downstream of the metabolic network in the DXP and MVA pathways (DXP, GPP, FPP, IPP + DMAPP). The generated time-series extracellular and intracellular metabolomics data from the engineered wild-type EcoCTs03 strain after IPTG induction was used to develop the dynamic model.

Dynamic model development with time-series metabolomics data and ^{13}C -tracer studies

The developed dynamic model utilised the time-series metabolomics data from the engineered wild-type EcoCTs03 strain after IPTG induction. To simplify the dynamic model, the rate law for most of the biochemical reactions in the metabolic network were designated as Michaelis–Menten reactions with metabolite inhibition accounted for according to literature^{62,63}. The metabolic pathway involving the PP pathway was also simplified by lumping the reactions producing isomers Ru5P, R5P, and X5P as the time-series data of these isomers were quantified as a pool on the LC-TOF-MS (liquid chromatography-time of flight-mass spectrometer)⁶⁴.

During the initial development of the limonene metabolic model (called Model A), the network topology was created (Supplementary Fig. 2) and the parameters of all dynamic equations used in Model A were fitted with the time-series measured metabolite concentrations using the genetic algorithm in COPASI⁶⁵ (Supplementary Fig. 3). Figure 3A shows that there

was poor fitting of extracellular limonene (LIMex) data with the simulated LIMex concentrations with time. In our previous work on modelling of the TLR (Toll-like receptor) 4, TNF (tumour necrosis factor), and TRAIL (TNF-related apoptosis-inducing ligand) signalling, we had a similar situation where the initial model created failed to fit experimental profiles^{35,36,66,67}. After carefully studying the local pathways where the model failed and using response rules, we were able to perform model iterations until satisfactory simulations were made. Here, we adopted a similar approach of performing model iterations for simulation improvements.

In Model A, notably, acetaldehyde (ACTLD) and ethanol (ETH) were found to accumulate (Fig. 3B, C). To reduce the accumulation of ACTLD and ETH, biochemical reactions downstream of these metabolites were further studied, where the reverse reaction of enzyme AtoB (acetyl-CoA acetyl-transferase) was found to be high. The reverse reaction of AtoB refers to the formation of AcCoA (acetyl coenzyme A) from AtAcCoA (acetoacetyl-CoA) and CoA (coenzyme A). Working from this reaction to the downstream reactions ensured that there was sufficient flux in the MVA pathway flowing towards LIMex. During model development, the approach of pulling the accumulated flux away from ACTLD and ETH towards the MVA pathway helped to improve the model further, resulting in a better refined Model B with improved fitting for LIMex (Fig. 3D). (The simulations for all measured metabolites for Model B can be found in Supplementary Fig. 4). Nevertheless, Model B had poor fitting for the simulations of the pooled metabolites GAP + DHAP (Fig. 3E). For both Models A and B, there was no initial inclusion of the ED pathway to the network topology (Supplementary Fig. 2) as this pathway seems to use gluconate instead of glucose^{12,13}. However, some studies have shown that *E. coli* can involve the EMP and ED pathways to metabolise glucose⁶⁸ while the isoprenoid synthesis pathways DXP and MVA⁶⁹ can result in the biosynthesis of limonene.

We next checked this, through the parallel use of two different glucose tracers, [1- ^{13}C]-glucose and [4- ^{13}C]-glucose, supplied in the cell culture

media, where ^{13}C atoms can be incorporated into the metabolite IPP (or isomer DMAPP) when glycolysis occurs via the EMP pathway using the $[1-^{13}\text{C}]$ -glucose, and for the $[4-^{13}\text{C}]$ -glucose tracer its incorporation into the isoprenoids (IPP and DMAPP) occurs via the DXP pathway⁶⁰. Since two IPP or DMAPP units are required for the biosynthesis of limonene, it is possible to determine molecular isotopologues having ^{13}C atoms incorporated through the analysis of the mass spectra of limonene obtained from the GC-MS. Each m/z in the mass distribution vector (MDV) relates to a unique number of ^{13}C -labelled IPPs integrated into limonene. By utilising $[1-^{13}\text{C}]$ -glucose as a positional label, there can be zero, one, two, or three ^{13}C atoms incorporated into IPP and DMAPP⁶⁰. A map describing the ^{13}C incorporation into IPP/DMAPP by parallel $[1-^{13}\text{C}]$ -glucose and $[4-^{13}\text{C}]$ -glucose labelling cultivation has been previously published⁶⁰. Supplementary Fig. 5 displays the MDV of limonene after correction for naturally occurring isotopes (MDV*). The interpretation of the MDV* of limonene needs the consideration of different labelling scenarios as predicted (Supplementary Fig. 6). As a result, conversion fraction $\frac{1}{2}$ was included as a multiplier for the $M + 1$ fraction while $\frac{3}{4}$ was included as multiplier for the $M + 2$ and $M + 3$ fractions (Supplementary Table 1). Using these conversion factors and the MDV* values, the calculation for the flux ratio for the EMP pathway was then determined, as previously described⁶⁰. The EMP pathway had a flux ratio of 0.945 ± 0.002 , while the ED pathway which is complementary to the contribution of the EMP pathway was found to have a flux ratio of 0.055 ± 0.002 . The results from the $[1-^{13}\text{C}]$ -glucose tracer studies showed that the EMP pathway was the main glycolysis pathway (94.5% contribution) while the ED pathway had a small contribution to glycolysis. Since the ED pathway was involved with a 5.5% contribution, it was included in the metabolic network of the next refined dynamic model C (Fig. 1). That is, the ED pathway, which results in the production of PYR and GAP + DHAP, was added to Model C. Notably, this helped in further improving the simulations of the pooled metabolites GAP + DHAP (Fig. 3F). Model C (Supplementary Fig. 7) also showed that the DXP pathway did contribute to limonene production, which was corroborated with experimental findings (Supplementary Fig. 8). The rate laws, rate law equations, and the parameter values of the final fitted model (Model C) can be found in Supplementary Table 2.

For cell culture cultivation with $[4-^{13}\text{C}]$ -glucose, at most only one ^{13}C -atom may be incorporated per IPP/DMAPP molecule⁶⁰. The DXP flux ratio was obtained through the MDV* value for this positional label and the

conversion fraction $\frac{1}{2}$ which was included as a multiplier for the $M + 1$ fraction (Supplementary Table 1). The DXP flux ratio was determined as 0.094 ± 0.001 , while the MVA flux ratio which also contributed to the isoprenoid flux was 0.906 ± 0.001 . This showed that the MVA pathway contributed significantly to isoprenoid biosynthesis and therefore limonene biosynthesis. There was also 9.4% contribution of the flux from the DXP pathway, indicating that this pathway is still involved in limonene biosynthesis. This contribution to limonene biosynthesis from the DXP pathway was also reflected in the final developed dynamic model (Model C), through the in silico knockout (KO) of the enzyme ISPH (hydroxymethylbutenyl 4-diphosphate reductase), where the simulation showed a slight reduction in limonene production as reflected in Supplementary Fig. 8 (the in silico KOs were performed by setting all the reaction parameters upstream of the model species, such as metabolites in this context, to zero^{35,70}).

The $[1-^{13}\text{C}]$ -glucose and $[4-^{13}\text{C}]$ -glucose tracer studies showed that the main route of carbon flux for limonene biosynthesis was through the EMP glycolytic pathway and MVA pathway. Based on the $[1-^{13}\text{C}]$ -glucose tracer studies, the EMP pathway was preferred over the ED pathway for glycolysis. This preference is likely due to the high expression of EMP pathway enzymes observed in *E. coli*⁶⁸. Furthermore, there was a possible synergy between the heterologous MVA and native DXP pathways, based on the results of the $[4-^{13}\text{C}]$ -glucose tracer studies, as both pathways contributed to limonene biosynthesis.

Utilising experimentally validated limonene model to identify key bottlenecks to enhance limonene yield

With model C satisfactorily recapitulating experimental dynamics for all measured metabolites concentrations and fluxes, next, we investigated, using this model, novel in silico targets to enhance limonene concentrations compared to the engineered wild-type EcoCTs03 strain. We performed in silico KOs for all enzymes and found ALDH KO, ADH KO, and LDH KO as the top targets which significantly increased the bioproduction of limonene. Since the enzymes ALDH and ADH are encoded by the same gene, *adhE*^{71,72}, the knockout of these enzymes was represented as the ALDH-ADH KO. Figure 4A shows simulations of extracellular limonene obtained from the in silico ALDH-ADH KO, where increased limonene yield was predicted. This was further qualitatively validated experimentally, where there was about eightfold more limonene produced 7 h post IPTG induction for the ALDH-

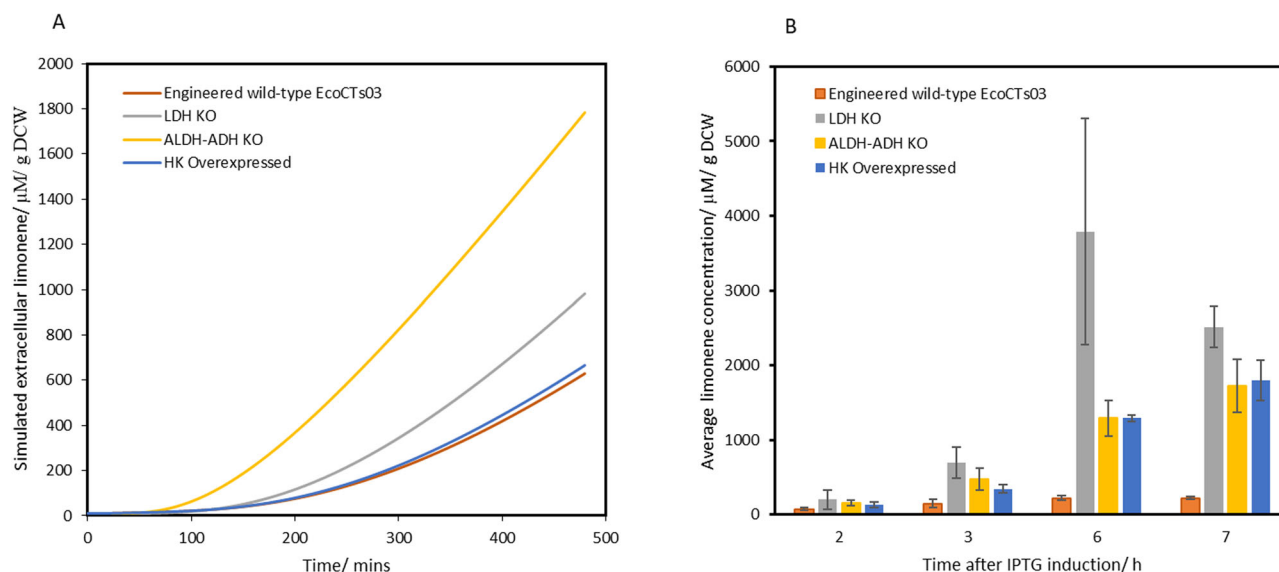


Fig. 4 | Extracellular limonene (LIMex) simulations of the in silico knockout (KO) of enzymes ALDH-ADH (aldehyde dehydrogenase-alcohol dehydrogenase) and LDH (lactate dehydrogenase), and overpressed HK (hexokinase) from the developed dynamic model with the corresponding experimental LIMex concentrations. A Compares the simulations of LIMex from the developed dynamic

model (Model C) for ALDH-ADH KO, LDH KO, HK overexpressed, and engineered wild-type EcoCTs03, while **B** shows the LIMex concentrations produced experimentally from the bacterial strains over time. Average values are shown for biological replicates (± 1 SD) in (B).

ADH KO strain compared to the engineered wild-type EcoCTs03 (Fig. 4B). For the LDH KO, the simulations showed an overall increase in limonene production (Fig. 4A) which corroborated with experimental results where there was significant (11-fold) increase in limonene produced 7 h post IPTG induction for the LDH knockout strain compared to the EcoCTs03 strain (Fig. 4B). The model also indicated that there was flux loss to the mixed fermentation pathway involving lactic acid and ethanol formation, as the LDH KO and ALDH-ADH KO simulations resulted in enhanced limonene yield.

We also performed a study on *in silico* overexpression by raising the relevant parameters (e.g. $V_{\max}$). Notably, when overexpressing the enzyme HK (*glk* gene) in *in silico*, the simulation of the extracellular concentration of limonene was found to increase (Fig. 4A) albeit less than that observed from the *in silico* ALDH-ADH KO and LDH KO. Experimentally, the HK overexpressed strain produced eightfold more limonene 7 h post IPTG induction compared to the engineered wild-type EcoCTs03 strain (Fig. 4B).

Discussion

Traditionally, increasing the yields of targets in microbial bio-factories have been carried out through various laborious experimental approaches such as manipulating genes from different pathways^{19,43}. However, such approaches may not be efficient in identifying mutant strains with the highest titres. Incorporating systems biology work, computational modelling approaches together with synthetic biology-based engineering, can help to steer metabolic engineering strategies to improve product yield through the identification of carbon flux-controlling fixtures in the highly integrated metabolic network. In this work, we generated intracellular and extracellular time-series metabolomics data from an engineered wild-type limonene-producing strain (EcoCTs03). The intracellular metabolites were quantified after fast filtration and rapid quenching in liquid nitrogen. The fast filtration and liquid nitrogen quenching methodology was preferred over cold methanol quenching as it minimises cell leakage from microbes and metabolite degradation whilst enabling the removal of contamination from extracellular sources^{73–76}. Quenching or the complete stopping of biochemical reactions involved in cell metabolism is vital for accurate quantitation of intracellular metabolites⁷⁷. Furthermore, since limonene is secreted extracellularly and not accumulated intracellularly due to its toxicity to the bacterial cells, a dodecane overlay in the cell culture was used to trap secreted limonene³ to ensure reliable quantitation due to limonene's volatility. LIMex was quantified from the dodecane layer and its concentration was found to generally increase through the sampling time points (Fig. 2B). LACex was only observed at the later time points from 6 to 8 h post IPTG induction (Fig. 2A) probably due to rapid increase in cell density resulting in some depletion in oxygen supply in the shake flasks. The fermentative conditions probably resulted in increased production of LACex from PYR^{78,79}. This also correlated with the drop in intracellular PYR observed over the same time points (Fig. 2D), where there was a diversion of carbon flux from PYR to LACex. The measured intracellular metabolites upstream of PYR, such as G6P, F6P, and DHAP + GAP, also had reduced concentrations during similar time points which correlated with more carbon flux used for limonene production and thereby increasing its yield.

The intracellular metabolites MVA and MVAP were not detected over the time points while MVAPP was only detected at a low concentration ($0.05 \pm 0.02 \mu\text{M/g DCW}$) at 8 h post IPTG induction. This probably shows that the carbon flux directed through the MVA pathway from MVA to MVAPP was efficiently converted during the sampling time points due to the lack of metabolite accumulation. The presence of intracellular FPP (Fig. 2D) showed that there was a loss of flux from GPP to LIMex production. Previous work has shown that expression of GPP synthase instead of the native *IspA* produced only GPP instead of both GPP and FPP, thereby redirecting the carbon flux towards GPP and preventing loss to FPP increased limonene titre³. The last two intracellular metabolites before conversion to LIMex, IPP + DMAPP and GPP were only detected at the later time points of 6–8 h post IPTG induction (Fig. 2C, D) which was correlated with the higher detection of LIMex (Fig. 2B). The increased

supply of these precursor metabolites (IPP + DMAPP and GPP) have also increased yield of target compounds in previous work^{7,80}.

Here we developed a dynamic model based on sound biochemical reaction equations and fitted with our *in-house* experimentally generated time-series metabolomics data. The dynamic model underwent few iterations (Models A and B) before the final developed model (Model C). In the initial Model A with the metabolic network topology of Supplementary Fig. 2, there was poor fitting of LIMex (Fig. 3A) due to the carbon flux accumulation found in the time course simulations of ACTLD and ETH (Fig. 3B, C). For the next model iteration, carbon flux was moved from ACTLD and ETH downstream towards LIMex, where Model B resulted in the improved fitting of LIMex (Fig. 3D). However, there was poor fitting of DHAP + GAP in Model B (Fig. 3E). For the final iteration, Model C incorporated the results from the [$1\text{-}^{13}\text{C}$]-glucose tracer studies where the ED pathway was added to the metabolic network topology (Fig. 1) which further improved the GAP + DHAP simulation (Fig. 3F). Model C also corroborated with [$4\text{-}^{13}\text{C}$]-glucose tracer studies, where most of the carbon flux went through the MVA pathway with minimal carbon flux entering the DXP pathway (Supplementary Fig. 8). The dynamic model was validated after showing *in silico* knockdowns of LDH and ALDH-ADH yielded enhanced limonene concentrations which was corroborated qualitatively with experimental results. Furthermore, *in silico* overexpression of HK from the dynamic model also improved limonene yield which was experimentally proven as well. The developed dynamic model was efficient in identifying enzyme knockouts and overexpressed enzyme that increased limonene yield.

From this systems biology work, most of the carbon flux flowed through the MVA pathway towards limonene biosynthesis, while a small amount of carbon flux went through the native DXP pathway. The small amount of carbon flux that went through the DXP pathway may probably be important for vitamin B6 synthesis, which is required for cellular processes and the biosynthesis and catabolism of amino acids^{81,82} as well as the production of the metabolite FPP, which is required for peptidoglycans formation for cell wall biosynthesis during early development⁸³ and in creating respiratory chain components^{84,85}.

The contributions of both the DXP and MVA pathways to limonene biosynthesis in the EcoCTs03 strain suggests possible synergistic interactions between the two isoprenoid routes. This hypothesised synergy between the DXP and MVA pathways has been observed in previous work in the bacterium *Rhodobacter sphaeroides*⁶⁰. The synergy of both pathways has also led to improved lycopene productivity in *E. coli*⁸⁶. In assessing the DXP and MVA pathways for IPP/DMAPP synthesis from glucose, the MVA pathway has greater energy efficiency due to the net increase in NAD(P)H reducing equivalents¹⁶ while the DXP pathway has greater carbon efficiency due to the requirement of only two GAP molecules for IPP/DMAPP synthesis compared to three GAP molecules by the MVA pathway¹⁶. The possible synergy between the DXP and MVA pathways could be due to the compatible nature of both pathways through co-factors, as the synthesis of one IPP/DMAPP from glucose through the MVA pathway results in a net of four NAD(P)H reducing equivalents, which could in turn generate ATP via respiration⁸⁶. Even though the synthesis of the MVA molecule through glycolysis, pyruvate dehydrogenase, and the upper MVA pathway results in NAD(P)H and ATP surplus, the lower MVA pathway results in no NAD(P)H gained and instead utilises ATP for the conversion of the MVA molecule to IPP/DMAPP. Therefore, the lower MVA pathway cannot provide the reducing equivalent demand and ATP to the DXP pathway, which probably results in a diminished flux going through the DXP pathway⁸⁶. The pairing of the complementary reducing power and ATP requirement therefore probably portrays a crucial part in the synergy of the dual pathway in limonene biosynthesis.

The presence of the ED pathway in EcoCTs03 was rather interesting as this pathway avoids the thermodynamic bottlenecks of fructose-1,6-bisphosphate aldolase and triose-phosphate isomerase faced by the EMP pathway¹¹, and requires substantially less enzymatic proteins than the EMP

pathway⁶⁸. However, the ED pathway, theoretically, results in one ATP produced compared to two ATP molecules produced by the EMP pathway (per molecule of glucose). In EcoCTs03, ATP was required for various reactions in the MVA pathway, for instance, in the biochemical reaction from metabolite MVA (mevalonate) to MVAP (5-phosphomevalonate), MVAP to MVAPP (5-diphosphomevalonate) and MVAPP to IPP/DMAPP (Fig. 1). In previous works, ED upregulation were found to enhance isoprenoid yields by increasing precursor metabolites^{87–89} and alleviating oxidative stress^{90,91}. However, less ATP produced through this pathway would make it difficult to improve limonene yield in this work. The ATP concentrations at various time points were not determined due to their concentrations being low, below the detection limit of the LC-TOF. The EMP pathway was much more dominant in EcoCTs03 probably due to the evolutionarily tuned EMP metabolic pathway found in *E. coli*, whereby EMP pathway enzymes have been observed to be highly expressed in the proteome⁶⁸.

The simulation of DXP pathway inactivation from the developed dynamic model for the EcoCTs03 strain, portrayed a slight reduction in limonene production, showing that the co-expression of both the DXP and MVA pathways helped to improve limonene production. This observation is further supported by experimental studies in *R. spaeroides*, where the co-expression of the DXP and MVA pathways improved amorphadiene titres by 1.2-fold, whilst the inactivation of the DXP pathway caused a threefold reduction in amorphadiene titres⁹². Moreover, in another experimental study in *R. spaeroides*, strains co-expressing the DXP and MVA pathways were found to produce the highest amorphadiene/biomass ratios compared to single-pathway strains cultivated in various conditions⁹³.

From the final refined model, in silico knockouts of LDH and ALDH-ADH showed favourably enhanced limonene yields. This was verified experimentally as eliminating the carbon flux losses by creating the ALDH-ADH and LDH KO strains helped to improve limonene yield between 8- and 11-folds. This was probably due to the availability of more carbon flux for limonene production with the elimination of competing pathways. With the LDH KO, there was less competition for pyruvate as a precursor and NADH as a co-factor for lactate production, ensuring more pyruvate availability for the MVA pathway. In a previous study, LDH and pyruvate formate lyase knockdowns were performed for *Enterobacter aerogenes* which improved NADH co-factor supply and 2,3-butanediol yield⁹⁴. For the ALDH-ADH KO, with the elimination of ethanol production, there was probably more metabolite acetyl coenzyme A (AcCoA) available for the MVA pathway and co-factor NADH. The final model also showed an improvement in limonene yield through in silico overexpression of HK. Limonene yield was also found to increase by eightfold through experimentally produced strains with HK overexpression, corroborating with the developed dynamic model. In another study, overexpressing the HK enzyme in engineered yeast also led to an increased production of the target compound, β -carotene⁹⁵. Although HK overexpression allowed improvement in limonene yield, there are further bottlenecks down the metabolic pathway, such as the loss of flux to mixed acid fermentation (lactic acid and ethanol formation) as shown through the improved limonene yield for the KO strains in Fig. 4A, B, and the possible reduction in ATP content due to its consumption in the increased phosphorylation of glucose by HK. Furthermore, the activity of HK could also be inhibited by the metabolite G6P (glucose-6-phosphate) produced in the HK enzymatic reaction which causes feedback inhibition^{96–98}. These factors could have prevented further improvements in limonene yield experimentally.

The identified knockout strains show that metabolic fluxes can be easily redirected through the deletion of competing pathways. Likewise, previous studies have shown that target compound yields can be improved through the elimination of by-products^{99,100}. Besides the deletion of competing pathways, product yield can also be enhanced through overexpression of target enzymes as shown in this study. The use of our model, therefore, removed the need for a trial-and-error-based experimental approach of creating a library of knockout and overexpressed strains to identify the best limonene-producing strains. The model also identified

three different strain manipulations which can enhance limonene production unlike previous work in cyanobacteria producing limonene which identified only one strain modification⁴². Further work on using these three identified strain manipulations as a combination in a single stable strain could aid in improving limonene yield even more. The supply of co-factors such as ATP or NADPH may also become limiting for the various biochemical reactions in the metabolic network. Fine-tuning the supply of such co-factors is another strategy that can aid in improving product yield^{42,94}.

Here we present a systems biology approach for limonene biosynthesis of engineered wild-type strain EcoCTs03 using time-series intracellular and extracellular metabolomics data, as well as incorporating ¹³C-tracer studies to develop a refined dynamic model. The model enabled the identification of key strain modifications in silico (ALDH-ADH KO, LDH KO, and HK overexpression) that would enhance the overall limonene yield. Based on the in silico results, both upstream and downstream bottlenecks in the metabolic network were identified, which in turn guided the construction of the identified modifications in the bacterial strains leading to significant limonene bioproduction. This strategy of incorporating a dynamic model into the metabolic engineering workflow to determine strain modifications resulting in improved target yields is more efficient than experimentally creating a large library of overexpressed or knockout strains to identify such high titre-producing modifications. Moreover, the possibility of incorporating the identified modifications into a single stable strain would aid in further improving target yield. After identifying key strain modifications with a developed dynamic model, further engineering strategies such as protein engineering and process optimisation may be utilised to improve product titres. Protein engineering could involve using more active homologous enzymes from alternative organisms which can improve product yield^{101–104}. Alternatively, targeted mutagenesis of key enzymes could also improve activity towards substrates¹⁰⁵.

Methods

All chemicals, standards, solvents, and media components were purchased from Sigma-Aldrich (St. Louis, MO), Fisher Scientific (Fair Lawn, NJ, USA), or VWR (Radnor, PA, USA). Deionised water was filtered by Sartorius Arium Pro VF Type 1 water system (18.2 M Ω , 0.2 μ m).

Strain construction

The engineered wild-type strain producing limonene, EcoCTs03, was produced by transforming pJBEI-6409 plasmid, which was a gift from Taek Soon Lee (Addgene plasmid # 47048; <http://n2t.net/addgene:47048>; RRID:Addgene_47048)³ through heat shock in *E. coli* K-12 MG1655. pJBEI-6409 is a p15A plasmid expressing limonene from acetyl-CoA via geranyl pyrophosphate (GPP) and was used as-is in this study. CRISPR Cas9-assisted recombineering method as described previously¹⁰⁶ was used to delete *ldhA* (UniProt P52643) and *adhE* (UniProt P0A9Q7) genes from the *E. coli* MG1655 strain, resulting in LDH and ALDH-ADH knockouts, respectively. The *ldhA* and *adhE* regions of three successfully knockout strains were PCR amplified and thereafter sequenced to confirm gene deletion (Supplementary Fig. 9). One colony each was selected and stored at -80°C in 40% glycerol prior to further use. To produce limonene, the pJBEI-6409 plasmid was subjected to heat shock in Δ *ldhA* and Δ *adhE* strains. The Mix & Go! *E. coli* Transformation Kit (Zymo Chem) was utilised following the manufacturer's instructions to prepare competent cells using the *ldhA* and *adhE* deletion strains. The pJBEI-6409 transformed strains (EcoCTs03, LDH KO, and ALDH-ADH KO) were placed onto individual chloramphenicol selective Luria-Bertani agar plates and left overnight at 37°C .

In this study, the plasmids used for the expression of the *glk* enzyme were derived from the previously published plasmids pJBEI-6409 and pTrc-trGPPS(CO)-LS (pJBEI-3101), which were obtained from Addgene³. While pJBEI-3101 is a ColE1 plasmid expressing limonene from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) via GPP,

Table 1 | Bacterial strains used in this study

Name	Description	Reference	Remarks
<i>E. coli</i> DH5a (ECOS 101)	endA1 recA1 relA1 deoR gyrA96 nupG hsdR17(rk-, mk +) supE44 thi-1 Δ(lacZYA-argF)U169 Φ80 lacZΔM15 λ- F-	Yeastern Biotech	Used for cloning.
<i>E. coli</i> K-12 MG1655	F- lambda- <i>ilvG</i> - <i>rfb</i> -50 <i>rph</i> -1	ATCC 700926	Used for limonene production.
EcoCTs03	<i>E. coli</i> K-12 MG1655 strain transformed with pJBEI-6409	This study	Strain for limonene production.
HK overexpressed	<i>E. coli</i> K-12 MG1655 strain transformed with pJBEI-6409 and EcoCTs-CPMS5- <i>glk</i>	This study	Strain for limonene production with glucokinase overexpression.
<i>E. coli</i> K-12 MG1655 Δ <i>ldhA</i>	<i>E. coli</i> K-12 MG1655 with deletion of <i>ldhA</i> gene	This study	Used for limonene production.
<i>E. coli</i> K-12 MG1655 Δ <i>adhE</i>	<i>E. coli</i> K-12 MG1655 with deletion of <i>adhE</i> gene	This study	Used for limonene production.
LDH KO	<i>E. coli</i> K-12 MG1655 Δ <i>ldhA</i> strain transformed with pJBEI-6409	This study	Strain for limonene production.
ALDH-ADH KO	<i>E. coli</i> K-12 MG1655 Δ <i>adhE</i> strain transformed with pJBEI-6409	This study	Strain for limonene production.

Table 2 | Plasmids used in this study

Name	Origin of replication	Antibiotic marker	Description	Reference
pJBEI-6409	p15A	Chloramphenicol	Limonene from acetyl-CoA via GPP. Plasmid map- UV5prom-AtoB-saHMGS-saHMGR-T1ter-pTrc-MK-PMK-PMD-idi-B0015term-pTrc-trGPPS-tLS-p15Aori-camR-lacI	pJBEI-6409 ³
pJBEI-3101	ColE1	Ampicillin	Limonene from IPP and DMAPP via GPP. Plasmid map-pTrc-trGPPS-tLS-t1ter-ampR-ColE1-lacI	pJBEI-3101 ³
EcoCTs-CPMS5	ColE1	Ampicillin	Plasmid for overexpression of limonene synthase. pJBEI-3101-del-trGPPS.	This study
EcoCTs-CPMS5- <i>glk</i>	ColE1	Ampicillin	Plasmid for overexpression of glucokinase	This study
pCas-V5	repA101	Kanamycin	Plasmid to express <i>cas9</i> under its native promoter, λRed from pBAD promoter.	106
pTargetT-pJ23119_ <i>adhE</i> -p- <i>sacB</i>	pMB1	Spectinomycin	Plasmid expresses gRNA, <i>SacB</i> under sucrose inducible promoter and harbours the homologous arm.	106
pTargetT-pJ23119_ <i>ldhA</i> -p- <i>sacB</i>	pMB1	Spectinomycin	Plasmid expresses gRNA, <i>SacB</i> under sucrose inducible promoter and harbours the homologous arm.	106

it was modified to express the selected enzyme of interest. A list of strains and primers can be found in Tables 1 and 2, respectively. The gene of interest was *glk* (Uniprot P0A6V8).

EcoCTs-CPMS5 was created by restriction-free cloning^{107,108}. Briefly, the plasmid backbone was amplified by PCR using iProof™ High-Fidelity DNA Polymerase (Bio-Rad), in which overlapping overhangs were created by primer design. The resulting DNA was treated with DpnI (New England Labs) overnight at 37 °C. From the DpnI-treated PCR product, 1 μL was transformed by heat shock in *E. coli* DH5a, plated on an ampicillin selective Luria-Bertani agar plate and incubated overnight at 37 °C. Positive colonies were identified by colony PCR using PCR BIO Taq mix red (PCR Biosystems), and were subsequently inoculated overnight in 2 mL Luria-Bertani supplemented with appropriate antibiotic at 37 °C, 300 rpm. Cultures were spun down, where the plasmid was extracted using E.Z.N.A.® Plasmid Mini Kit I (Omega Bio-tek) and sequenced using Sanger sequencing (Bio Basic) to confirm the presence of the deletion. Finally, plasmids were confirmed by performing restriction enzyme digestion (RE) and compared with in silico agarose gel simulation on SnapGene 7.0.2.

The gene of interest was cloned in EcoCTs-CPMS5 by PCR-amplifying the backbone of the plasmid without limonene synthase (LS), as well as the genes of interest from *E. coli* K-12 MG1655 genomic DNA. Primers contained overhangs such that the backbone and inserts shared a 15 bp overlap region at the site of integration. Expected band sizes were confirmed on an agarose gel and subsequently treated with DpnI overnight at 37 °C. DpnI-treated PCR products were purified using the E.Z.N.A.® gel extraction kit (Omega Bio-tek). Each of purified DNA backbone (50–100 ng) and inserts were added to a PCR tube with In-Fusion Snap Assembly Master Mix (Takara Bio). The mixture was incubated at 50 °C for 15 min, and 1 μL was transformed by heat shock in *E. coli* DH5a, plated on an ampicillin selective Luria-Bertani agar plate and incubated overnight at 37 °C. Colonies were analysed by colony PCR, amplified, extracted, sequenced, and RE digested as described above.

For expression, plasmids were transformed by heat shock in *E. coli* K-12 MG1655. Competent cells were prepared using the Mix & Go! *E. coli* Transformation Kit (Zymo Chem) following the manufacturer’s specifications. For the expression experiment, pJBEI-6409 was co-transformed with in EcoCTs-CPMS5-GOI (gene of interest, *glk*), plated on ampicillin and chloramphenicol selective LB agar plates and incubated overnight at 37 °C.

Growth conditions

From the prepared *E. coli* strains which were placed onto plates, one colony from each plate was selected after overnight incubation at 37 °C. Each colony was inoculated into 5 mL Luria-Bertani medium (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl) with 30 μg/mL chloramphenicol and left overnight in a shaking incubator at 220 rpm and 37 °C. After which, cell pellets were washed and resuspended in 50 mL M9 medium (12.7 g/L Na₂HPO₄·7H₂O, 3.1 g/L KH₂PO₄, 1 g/L NH₄Cl, 0.5 g/L NaCl, 0.25 g/L MgSO₄·7H₂O, 15 mg/L CaCl₂·2H₂O, 8.1 mg/L FeCl₃, 0.89 mg/L MnCl₂·4H₂O, 1.7 mg/L ZnCl₂, 0.34 mg/L CuCl₂, 0.6 mg/L CoCl₂·6H₂O, 0.51 mg/L Na₂MoO₄) adapted from a previous study¹⁰⁹ with 10 g/L glucose and left overnight in a shaking incubator at 30 °C and 220 rpm. Each strain was kept at –80 °C in 40% glycerol. For each glycerol-stored strain, 100 μL was added to 50 mL M9 medium in 250-mL flasks forming pre-cultures and left overnight at 220 rpm and 30 °C. For the collection of metabolomics data at various time points, numerous cell culture flasks were prepared for duplicate biological samples for each time-series collection. Through the addition of 100 μL pre-cultures of each strain to 50 mL M9 medium in 250-mL flasks, cell cultures of each strain were prepared and kept overnight in a shaking incubator at 220 rpm and 30 °C. Isopropyl β-d-1-thiogalactopyranoside (IPTG) was added resulting in a final concentration of 25 μM when cell cultures reached an optical density of 1 at 600 nm. To enable trapping of secreted limonene, dodecane overlays of 5 mL were added to cell cultures³ and kept at 30 °C and 220 rpm in a shaking incubator. To obtain the concentrations of the intracellular and extracellular metabolites, and secreted

limonene for dynamic modelling of the EcoCTs03 strain, prepared cell culture flasks were sacrificed in duplicates at time points 2 h, 3 h, 6 h, 7 h and 8 h post IPTG induction.

Metabolomics extraction and analysis

For the analysis of secreted limonene and extracellular metabolites, cell cultures of 50 mL EcoCTs03 with dodecane overlay were sacrificed at time points 0 h, 2 h, 3 h, 6 h, 7 h, and 8 h post IPTG induction in duplicates and centrifuged at 3000 rpm for 10 min. For the mutant strains (HK overexpressed, ALDH-ADH knockout, and LDH knockout), only secreted limonene concentrations were analysed by sacrificing shake flasks in duplicates at time points 2 h, 3 h, 6 h, and 7 h post IPTG induction. In addition, flasks with EcoCTs03 and HK overexpressed were analysed for protein expression. Briefly, the optical density at the final time point was measured and a normalised amount of cells was extracted ($V = 250/\text{OD } \mu\text{L}$). It was then spun down for 2 min at room temperature at 14,000 rpm. The media was discarded, where the pellets were resuspended in 50 μL B-PER, and vortexed for 10 min at room temperature. The resulting lysate was loaded on SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis), where the overexpression of HK was confirmed (Supplementary Fig. 10). The dodecane overlay which trapped the secreted limonene was removed from each sample and stored at -80°C prior to analysis. Diluted limonene extracts in ethyl acetate were run on an Agilent 7890B gas chromatography-mass spectrometry (GC-MS) system equipped with a DB-5ms column. A 10:1 split ratio and 10 mL/min split flow was utilised for each run with 1 μL of sample injected. The GC oven was held at 40°C for 3 min, with a temperature gradient of $10^\circ\text{C}/\text{min}$ to 100°C followed by a $60^\circ\text{C}/\text{min}$ ramp to 220°C , which was held for 2 min. The temperatures of the injector and MS transfer line were set at 250°C and 280°C , respectively. Selected ion-monitoring (SIM) mode was utilised for the MS with ions of m/z 136, 68 and 93 monitored, which represented the molecular ion and top two abundant fragmental ions of limonene. Calibration standards were prepared for limonene in ethyl acetate with concentrations ranging from 0.05 $\mu\text{g}/\text{mL}$ to 10 $\mu\text{g}/\text{mL}$. Limonene concentrations from cell cultures were determined using Agilent Quantitative software.

After centrifugation, supernatants of the EcoCTs03 cell cultures were kept at -80°C before subjecting the extracellular metabolites to quantitative analysis. Aliquots of 1 mL from the thawed supernatant samples were filtered using polyamide filters and subjected to an Agilent 1200 high-performance liquid chromatography (HPLC) system equipped with a Bio-rad Aminex HPX-87H column ($300 \times 7.8 \text{ mm}$). The HPLC system was run together with a 1260 Infinity II Refractive Index Detector (RID). An isocratic gradient was run for 28 min with 0.01 N sulphuric acid, 0.6 mL/min flow rate, and 5 μL of sample injected. The HPLC column was at 35°C while the RID had positive polarity at 30°C . Concentrations of extracellular metabolites were determined from calibrations mixtures of the following range: glucose—0.5 to 80 g/L; acetic acid—0.125–80 g/L; lactic acid—0.125–8 g/L; ethanol—0.5–80 g/L. Cell pellets left after centrifugation were oven-dried before weighing.

For the analysis of intracellular metabolites, cell cultures of EcoCTs03 were sacrificed in duplicates at time points of 2 h, 3 h, 6 h, 7 h, and 8 h after IPTG induction, where 10 mL aliquots of cell cultures were subjected to fast filtration followed by rapid quenching in liquid nitrogen as described previously with modifications^{73,110}. Briefly, for fast filtration, 10 mL aliquots of each cell culture were filtered using 0.2 μm polyamide membrane filters (Sartorius, Goettingen, Germany) followed by washing with 5 mL of wash solution (1 g/L NH_4Cl , 0.5 g/L NaCl, 12.7 g/L $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 3.1 g/L KH_2PO_4). The filter membrane containing the cells was placed onto aluminum foil, folded, and quickly plunged into liquid nitrogen. The membrane filters encased in aluminum foil were stored at -80°C until the extraction of the intracellular metabolites. The intracellular metabolites were extracted by removing the aluminum foil and placing the membranes in 5 mL of a 4:4:2 mixture of methanol, acetonitrile, and water, and left to stand in an ice bath for 10 min. After which, membranes were vortexed for

1 min and subjected to sonication for 3 min thrice, where samples were placed in an ice bath for 1 min between each round of sonication. An internal standard mixture consisting of 10 $\mu\text{g}/\text{mL}$ thymolphthalein mono-phosphate (TMP) and 50 $\mu\text{g}/\text{mL}$ mevalonic acid-d3 (MVA-d3) in a solvent mixture of methanol: 10 mM ammonium hydroxide (7:3) was prepared. Extracts were spiked with 20 μL of the internal standard mixture after decanting into glass tubes. This was followed by subjection to a vacuum concentrator and reconstitution with 200 μL methanol: 10 mM ammonium hydroxide mixture (7:3) prior to filtration into glass vials for quantitative liquid chromatography-mass spectrometry analysis.

An Agilent 6230 time of flight-mass spectrometer (TOF-MS) coupled with a Dual Agilent Jet Stream (AJS) ion source was used together with an Agilent ultra-performance liquid chromatography (UPLC) 1290 system. A VanGuard pre-column ($2.1 \times 5 \text{ mm}$) was utilised with a Waters Acquity UPLC BEH C18 column ($2.1 \times 150 \text{ mm}$, 1.7 μm). The chromatographic method utilised was modified from previous work^{73,110}. Samples of 2 μL were injected into the system with mobile phase A consisting of 5 mM ammonium formate in water (pH 9.5) and mobile phase B consisting of 5 mM ammonium formate (pH 9.5) in acetonitrile: water (9:1). The solvent gradient utilised for each run started with a flow rate of 0.1 mL/min and 100% mobile phase A held from 0 to 3.5 min followed by 100% mobile phase B at 12 min which was then held for 8 min and flow rate increased to 0.5 mL/min. The mobile phase was changed to 100% mobile phase A at 20 min for 5 min and kept for another 5 min. The column temperature was held at 35°C throughout the analysis. For the TOF, negative electrospray ionisation was utilised with the following conditions: Gas flow, 11 L/min; Gas temperature, 325°C ; Nebuliser pressure, 35 psi; Sheath gas flow, 11 L/min; Sheath gas temperature, 375°C ; Vcap voltage, 3500 V; Nozzle voltage, 500 V; Skimmer, 65; OctopoleRFPeak, 750; Scan rate, 2 spectra/s. The fragmentor voltage was altered during each 35 min sample run: 2–7.5 min, 140 V; 7.5–15 min, 100 V, 140 V and 150 V. The following flow diversions for the UPLC was used: 0–2 min to waste, 2–15 min to TOF-MS, and 15–35 min to waste. Methanol: 10 mM ammonium hydroxide (7:3) mixture was used for preparing standard mixtures used for intracellular metabolite quantitation. The intracellular metabolites determined were R5P + Ru5P + X5P (pool of ribose-5-phosphate, ribulose-5-phosphate, and xylulose-5-phosphate), F16BP (fructose-1,6-biphosphate), DHAP + GAP (pool of dihydroxyacetone phosphate and glyceraldehyde-3-phosphate), PYR (pyruvate), DXP (1-deoxy-D-xylulose-5-phosphate), MVA (mevalonate), MVAP (5-phosphomevalonate), MVAPP (5-pyrophosphomevalonate), IPP + DMAPP (pool of isopentenyl pyrophosphate and dimethylallyl pyrophosphate), FPP (farnesyl diphosphate), and GPP (geranyl diphosphate). Calibration mixtures were made up to 100 μL with 10 μL of internal standard mixture consisting of 50 $\mu\text{g}/\text{mL}$ MVA-d3 and 10 $\mu\text{g}/\text{mL}$ TMP. The following concentration range was used for the calibration mixtures: R5P (R5P + Ru5P + X5P pool)—0.04–10 $\mu\text{g}/\text{mL}$; F16BP—0.04–6 $\mu\text{g}/\text{mL}$; DHAP (DHAP + GAP pool)—0.04–5 $\mu\text{g}/\text{mL}$; PYR—0.04–1.5 $\mu\text{g}/\text{mL}$; DXP—0.04 to 10 $\mu\text{g}/\text{mL}$; MVA—0.04–10 $\mu\text{g}/\text{mL}$; MVAP—0.01–0.3 $\mu\text{g}/\text{mL}$; MVAPP—0.01–0.3 $\mu\text{g}/\text{mL}$; IPP (IPP + DMAPP pool)—0.05–1.5 $\mu\text{g}/\text{mL}$; FPP—0.05–1.5 $\mu\text{g}/\text{mL}$; GPP—0.05–1.5 $\mu\text{g}/\text{mL}$. Agilent Masshunter Workstation Quantitative Analysis for TOF was utilised for metabolite quantitation.

G6P (glucose-6-phosphate) and F6P (fructose-6-phosphate) were quantified by injecting 2 μL of samples and running the samples through the LC-TOF-MS system with a ZIC-HILIC column ($2.1 \times 100 \text{ mm}$, 3.5 μm). The solvent gradient started with a 10% mobile phase A with a flow rate of 0.5 mL/min at 0 min, 25% mobile phase A at 1.5 min, 35% mobile phase A at 1.8 min and held until 6 min, 10% mobile phase A at 6.5 min until 9.5 with flow rate of 0.6 mL/min, and 10% mobile phase with flow rate of 0.5 mL/min held until 10 min. The column temperature was constant at 35°C . For the TOF, the same conditions were utilised as described for the other intracellular metabolites with exception to the fragmentor voltage which was kept at 140 V from 0.1–6 min while the flow diversion for UPLC was 0–0.1 min to waste, 0.1–6 min to TOF-MS, and 6–10 min to waste. The following

concentration range was used for the calibration mixtures: G6P—0.04–10 µg/mL and F6P—0.04–6 µg/mL, with TMP as the internal standard. Metabolite quantitation was executed using the Agilent Masshunter Workstation Quantitative Analysis for TOF.

For experiments involving ^{13}C tracers, $[1-^{13}\text{C}]\text{glucose}$ (99 atom% ^{13}C) and $[4-^{13}\text{C}]\text{glucose}$ (99 atom% ^{13}C) tracers were used in parallel in independent 50 mL cell cultures of the engineered wild-type strain (EcoCTs03) instead of glucose as described in the previous section. The cell cultures were prepared in duplicates and sacrificed 24 h after IPTG induction for cell cultures containing either $[1-^{13}\text{C}]\text{glucose}$ or $[4-^{13}\text{C}]\text{glucose}$. The dodecane layers were removed after centrifugation and subjected to GC-MS analysis in scan mode for limonene, where the metabolic flux ratios for the EMP and ED pathway based on the $[1-^{13}\text{C}]\text{glucose}$ and the metabolic flux ratios for DXP and MVA pathway based on the $[4-^{13}\text{C}]\text{glucose}$ were determined based on previous work⁶⁰. The correction of limonene mass distribution vectors (MDVs) for natural ^{13}C and ^2H isotopes incorporation was executed as previously described¹¹¹, where the resulting corrected MDV (MDV*) was utilised for the calculation of the metabolic flux ratios.

Model construction

The dynamic model of limonene production by the *E. coli* strain EcoCTs03 was developed with the open-source and stand-alone program COPASI (build 260⁶¹) and it was modified from our recently published model⁶⁵. The dynamic model describes the carbon and energy metabolism of EcoCTs03 expressing the MVA pathway from the exponential growth phase after IPTG induction under aerobic conditions. The model contains 55 species and 56 reactions that consist of the pathways found in Fig. 1, such as EMP pathway, ED pathway, tricarboxylic acid cycle (TCA cycle), pentose phosphate pathway, acetate metabolism, MVA pathway and DXP pathway^{62,112,113}. For model simplicity, the model only comprises one compartment. Extracellular metabolites were denoted with an “-ex” suffix, for example “Glcex” or “ACEex” which denotes extracellular glucose and acetate respectively. Concentration units of metabolites in the model were in µmol/L/g dry cell weight (g DCW; reflected as µmol/l in the COPASI file), which was the same units utilised in the time-series metabolomics data. Normalising metabolite concentrations to g DCW allowed for comparison across the different time points after IPTG induction. All reactions were described using Michaelis–Menten rate law, except for enzymes PDH, LDH, PoxB and LAC transport reactions which were described by mass action instead. Supplementary Table 1 details the rate laws, rate law equations and the parameter values of the final fitted model.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information.

Code availability

The dynamic limonene COPASI model code is available at https://github.com/heeyanting93/CBioOmics_COPASI/tree/main/npj_model.

Received: 17 April 2024; Accepted: 19 September 2024;

Published online: 01 October 2024

References

- Gershenzon, J. & Dudareva, N. The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **3**, 408–414 (2007).
- Colby, S. M. et al. 4S-limonene synthase from the oil glands of spearmint (*Mentha spicata*). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. *J. Biol. Chem.* **268**, 23016–23024 (1993).
- Alonso-Gutierrez, J. et al. Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production. *Metab. Eng.* **19**, 33–41 (2013).
- Tracy, N. I. et al. Hydrogenated monoterpenes as diesel fuel additives. *Fuel* **88**, 2238–2240 (2009).
- Duetz, W. A. et al. Biotransformation of limonene by bacteria, fungi, yeasts, and plants. *Appl. Microbiol. Biotechnol.* **61**, 269–277 (2003).
- Ciriminna, R. et al. Limonene: a versatile chemical of the bioeconomy. *Chem. Commun.* **50**, 15288–15296 (2014).
- Carter, O. A., Peters, R. J. & Croteau, R. Monoterpene biosynthesis pathway construction in *Escherichia coli*. *Phytochemistry* **64**, 425–433 (2003).
- Reiling, K. K. et al. Mono and diterpene production in *Escherichia coli*. *Biotechnol. Bioeng.* **87**, 200–212 (2004).
- Willrodt, C. et al. Engineering the productivity of recombinant *Escherichia coli* for limonene formation from glycerol in minimal media. *Biotechnol. J.* **9**, 1000–1012 (2014).
- Yang, J. et al. Metabolic engineering of *Escherichia coli* for the biosynthesis of alpha-pinene. *Biotechnol. Biofuels* **6**, 60 (2013).
- Hollinshead, W. D. et al. Examining *Escherichia coli* glycolytic pathways, catabolite repression, and metabolite channeling using Δpfk mutants. *Biotechnol. Biofuels* **9**, 212 (2016).
- Eisenberg, R. C. & Dobrogosz, W. J. Gluconate metabolism in *Escherichia coli*. *J. Bacteriol.* **93**, 941–949 (1967).
- Peekhaus, N. & Conway, T. What's for dinner?: Entner–Doudoroff metabolism in *Escherichia coli*. *J. Bacteriol.* **180**, 3495–3502 (1998).
- Zhou, K., Edgar, S. & Stephanopoulos, G. Engineering microbes to synthesize plant isoprenoids. *Methods Enzymol.* **575**, 225–245 (2016).
- Ward, V. C., Chatzivasileiou, A. O. & Stephanopoulos, G. Metabolic engineering of *Escherichia coli* for the production of isoprenoids. *FEMS Microbiol. Lett.* **365**, fny079 (2018).
- Gruchattka, E. et al. In silico profiling of *Escherichia coli* and *Saccharomyces cerevisiae* as terpenoid factories. *Microb. Cell Fact.* **12**, 84 (2013).
- Diner, B. A. et al. Synthesis of heterologous mevalonic acid pathway enzymes in *Clostridium ljungdahlii* for the conversion of fructose and of syngas to mevalonate and isoprene. *Appl. Environ. Microbiol.* **84**, e01723–17 (2018).
- Sun, C., Theodoropoulos, C. & Scrutton, N. S. Techno-economic assessment of microbial limonene production. *Bioresour. Technol.* **300**, 122666 (2020).
- Ren, Y. et al. Microbial production of limonene and its derivatives: achievements and perspectives. *Biotechnol. Adv.* **44**, 107628 (2020).
- Martin, V. J. J. et al. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **21**, 796–802 (2003).
- Li, Q. et al. Balanced activation of IspG and IspH to eliminate MEP intermediate accumulation and improve isoprenoids production in *Escherichia coli*. *Metab. Eng.* **44**, 13–21 (2017).
- Liao, P. et al. The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnol. Adv.* **34**, 697–713 (2016).
- Shin, J., South, E. J. & Dunlop, M. J. Transcriptional tuning of mevalonate pathway enzymes to identify the impact on limonene production in *Escherichia coli*. *ACS Omega* **7**, 18331–18338 (2022).
- Wu, J. et al. Systematic optimization of limonene production in engineered *Escherichia coli*. *J. Agric Food Chem.* **67**, 7087–7097 (2019).
- Rohmer, M. et al. Glyceroldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative non-mevalonate pathway for terpenoid biosynthesis. *J. Am. Chem. Soc.* **118**, 2564–2566 (1996).
- Helmy, M., Smith, D. & Selvarajoo, K. Systems biology approaches integrated with artificial intelligence for optimized metabolic engineering. *Metab. Eng. Commun.* **11**, e00149 (2020).

27. Strutz, J. et al. Metabolic kinetic modeling provides insight into complex biological questions, but hurdles remain. *Curr. Opin. Biotechnol.* **59**, 24–30 (2019).
28. Kim, W. J., Kim, H. U. & Lee, S. Y. Current state and applications of microbial genome-scale metabolic models. *Curr. Opin. Syst. Biol.* **2**, 10–18 (2017).
29. Schellenberger, J. et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **6**, 1290–1307 (2011).
30. Heirendt, L. et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14**, 639–702 (2019).
31. Saa, P. A. & Nielsen, L. K. Formulation, construction and analysis of kinetic models of metabolism: A review of modelling frameworks. *Biotechnol. Adv.* **35**, 981–1003 (2017).
32. Choudhury, S. et al. Reconstructing kinetic models for dynamical studies of metabolism using generative adversarial networks. *Nat. Mach. Intell.* **4**, 710–719 (2022).
33. Kim, O. D., Rocha, M. & Maia, P. A review of dynamic modeling approaches and their application in computational strain optimization for metabolic engineering. *Front. Microbiol.* **9**, 1690 (2018).
34. Rapoport, T. A., Heinrich, R. & Rapoport, S. M. The regulatory principles of glycolysis in erythrocytes in vivo and in vitro. A minimal comprehensive model describing steady states, quasi-steady states and time-dependent processes. *Biochem J.* **154**, 449–469 (1976).
35. Selvarajoo, K. et al. Signaling flux redistribution at toll-like receptor pathway junctions. *PLoS ONE* **3**, e3430 (2008).
36. Hayashi, K. et al. A systems biology approach to suppress TNF-induced proinflammatory gene expressions. *Cell Commun. Signal* **11**, 84 (2013).
37. Selvarajoo, K. A systems biology approach to overcome TRAIL resistance in cancer treatment. *Prog. Biophys. Mol. Biol.* **128**, 142–154 (2017).
38. Selvarajoo, K. Complexity of biochemical and genetic responses reduced using simple theoretical models. *Methods Mol. Biol.* **1702**, 171–201 (2018).
39. Ryu, D. D. & Park, S. H. Genetically structured kinetic model for gene product and application of gene switching system to fermentation process control. *Ann. N. Y. Acad. Sci.* **506**, 396–405 (1987).
40. Hargrove, J. L., Hulse, M. G. & Beale, E. G. The kinetics of mammalian gene expression. *Bioessays* **13**, 667–674 (1991).
41. Thornburg, Z. R. et al. Kinetic modeling of the genetic information processes in a minimal cell. *Front. Mol. Biosci.* **6**, 130 (2019).
42. Wang, X. et al. Enhanced limonene production in cyanobacteria reveals photosynthesis limitations. *Proc. Natl. Acad. Sci. USA* **113**, 14225–14230 (2016).
43. Halfmann, C., Gu, L. & Zhou, R. Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO₂ and H₂O. *Green. Chem.* **16**, 3175–3185 (2014).
44. Breitling, R. What is systems biology? *Front. Physiol.* **1**, 9 (2010).
45. Likić, V. A. et al. Systems biology: the next frontier for bioinformatics. *Adv. Bioinforma.* **2010**, 268925 (2010).
46. Fong, S. S. et al. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.* **91**, 643–648 (2005).
47. Burgard, A. P., Pharkya, P. & Maranas, C. D. OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657 (2003).
48. Pharkya, P., Burgard, A. P. & Maranas, C. D. OptStrain: a computational framework for redesign of microbial production systems. *Genome Res.* **14**, 2367–2376 (2004).
49. Kim, J. & Reed, J. L. OptORF: optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst. Biol.* **4**, 53 (2010).
50. Rocha, I. et al. OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Syst. Biol.* **4**, 45 (2010).
51. Ranganathan, S., Suthers, P. F. & Maranas, C. D. OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Comput. Biol.* **6**, e1000744 (2010).
52. Chowdhury, A., Zomorodi, A. R. & Maranas, C. D. k-OptForce: integrating kinetics with flux balance analysis for strain design. *PLoS Comput. Biol.* **10**, e1003487 (2014).
53. Zhuang, K. et al. Dynamic strain scanning optimization: an efficient strain design strategy for balanced yield, titer, and productivity. DySScO strategy for strain design. *BMC Biotechnol.* **13**, 8 (2013).
54. Fong, S. S. Computational approaches to metabolic engineering utilizing systems biology and synthetic biology. *Comput. Struct. Biotechnol. J.* **11**, 28–34 (2014).
55. Chung, C. H. et al. Next-generation genome-scale metabolic modeling through integration of regulatory mechanisms. *Metabolites* **11**, 606 (2021).
56. Khana, D. B., Callaghan, M. M. & Amador-Noguez, D. Novel computational and experimental approaches for investigating the thermodynamics of metabolic networks. *Curr. Opin. Microbiol.* **66**, 21–31 (2022).
57. Vivek-Ananth, R. P. & Samal, A. Advances in the integration of transcriptional regulatory information into genome-scale metabolic models. *Biosystems* **147**, 1–10 (2016).
58. Selvarajoo, K. Parameter-less approaches for interpreting dynamic cellular response. *J. Biol. Eng.* **8**, 23 (2014).
59. Yeo, H. C. & Selvarajoo, K. Machine learning alternative to systems biology should not solely depend on data. *Brief. Bioinforma.* **23**, bbac436 (2022).
60. Orsi, E. et al. Metabolic flux ratio analysis by parallel ¹³C labeling of isoprenoid biosynthesis in *Rhodobacter sphaeroides*. *Metab. Eng.* **57**, 228–238 (2020).
61. Hoops, S. et al. COPASI—a complex pathway simulator. *Bioinformatics* **22**, 3067–3074 (2006).
62. Boecker, S. et al. Deciphering the physiological response of *Escherichia coli* under high ATP demand. *Mol. Syst. Biol.* **17**, e10504 (2021).
63. Costello, Z. & Martin, H. G. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ Syst. Biol. Appl.* **4**, 19 (2018).
64. Schuster, S., Fell, D. A. & Dandekar, T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.* **18**, 326–332 (2000).
65. Khanijou, J. K., Hee, Y. T. & Selvarajoo, K. Identifying key in silico knockout for enhancement of limonene yield through dynamic metabolic modelling. In *Systems Biology* (ed. Bizzarri, M.) 3–19 (Springer US, 2024).
66. Piras, V. et al. Enhancing apoptosis in TRAIL-resistant cancer cells using fundamental response rules. *Sci. Rep.* **1**, 144 (2011).
67. Hayashi, K. et al. Systems biology strategy reveals PKC- δ is key for sensitizing TRAIL-resistant human fibrosarcoma. *Front. Immunol.* **5**, 659 (2015).
68. Flamholz, A. et al. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci. USA* **110**, 10039–10044 (2013).
69. Kirby, J. & Keasling, J. D. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu. Rev. Plant Biol.* **60**, 335–355 (2009).

70. Selvarajoo, K. Discovering differential activation machinery of the Toll-like receptor 4 signaling pathways in MyD88 knockouts. *FEBS Lett.* **580**, 1457–1464 (2006).
71. Bertsch, J. et al. A novel route for ethanol oxidation in the acetogenic bacterium *Acetobacterium woodii*: the acetaldehyde/ethanol dehydrogenase pathway. *Environ. Microbiol.* **18**, 2913–2922 (2016).
72. Kim, G. et al. Aldehyde-alcohol dehydrogenase undergoes structural transition to form extended spiroosomes for substrate channeling. *Commun. Biol.* **3**, 298 (2020).
73. Castaño-Cerezo, S. et al. Functional analysis of isoprenoid precursors biosynthesis by quantitative metabolomics and isotopologue profiling. *Metabolomics* **15**, 115 (2019).
74. Heuillet, M. et al. Methodology for the validation of isotopic analyses by mass spectrometry in stable-isotope labeling experiments. *Anal. Chem.* **90**, 1852–1860 (2018).
75. da Luz, J. A., Hans, E. & Zeng, A.-P. Automated fast filtration and on-filter quenching improve the intracellular metabolite analysis of microorganisms. *Eng. Life Sci.* **14**, 135–142 (2014).
76. van Gulik, W. M. Fast sampling for quantitative microbial metabolomics. *Curr. Opin. Biotechnol.* **21**, 27–34 (2010).
77. Pinu, F. R., Villas-Boas, S. G. & Aggio, R. Analysis of intracellular metabolites from microorganisms: quenching and extraction protocols. *Metabolites* **7**, 53 (2017).
78. Clark, D. P. The fermentation pathways of *Escherichia coli*. *FEMS Microbiol. Rev.* **5**, 223–234 (1989).
79. Förster, A. H. & Gescher, J. Metabolic engineering of *Escherichia coli* for production of mixed-acid fermentation end products. *Front. Bioeng. Biotechnol.* **2**, 16 (2014).
80. Lin, P.-C., Zhang, F. & Pakrasi, H. B. Enhanced limonene production in a fast-growing cyanobacterium through combinatorial metabolic engineering. *Metab. Eng. Commun.* **12**, e00164 (2021).
81. Commichau, F. M. et al. Overexpression of a non-native deoxyxylulose-dependent vitamin B6 pathway in *Bacillus subtilis* for the production of pyridoxine. *Metab. Eng.* **25**, 38–49 (2014).
82. Mooney, S. et al. Vitamin B6: a long known compound of surprising complexity. *Molecules* **14**, 329–351 (2009).
83. Apfel, C. M. et al. Use of genomics to identify bacterial undecaprenyl pyrophosphate synthetase: cloning, expression, and characterization of the essential *uppS* gene. *J. Bacteriol.* **181**, 483–492 (1999).
84. Okada, K. et al. The *ispB* gene encoding octaprenyl diphosphate synthase is essential for growth of *Escherichia coli*. *J. Bacteriol.* **179**, 3058–3060 (1997).
85. Saiki, K., Mogi, T. & Anraku, Y. Heme O biosynthesis in *Escherichia coli*: the *cyo* gene in the cytochrome BO operon encodes a protoheme IX farnesyltransferase. *Biochem. Biophys. Res. Commun.* **189**, 1491–1497 (1992).
86. Yang, C. et al. Synergy between methylerythritol phosphate pathway and mevalonate pathway for isoprene production in *Escherichia coli*. *Metab. Eng.* **37**, 79–91 (2016).
87. Ng, C. Y. et al. Rational design of a synthetic Entner–Doudoroff pathway for improved and controllable NADPH regeneration. *Metab. Eng.* **29**, 86–96 (2015).
88. Li, C. et al. Modification of targets related to the Entner–Doudoroff/pentose phosphate pathway route for methyl-d-erythritol 4-phosphate-dependent carotenoid biosynthesis in *Escherichia coli*. *Microb. Cell Fact.* **14**, 117 (2015).
89. Liu, H. et al. MEP pathway-mediated isopentenol production in metabolically engineered *Escherichia coli*. *Microb. Cell Fact.* **13**, 135 (2014).
90. Chavarría, M. et al. The Entner–Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. *Environ. Microbiol.* **15**, 1772–1785 (2013).
91. Klingner, A. et al. Large-Scale ¹³C flux profiling reveals conservation of the Entner–Doudoroff pathway as a glycolytic strategy among marine bacteria that use glucose. *Appl Environ. Microbiol.* **81**, 2408–2422 (2015).
92. Orsi, E. et al. Functional replacement of isoprenoid pathways in *Rhodobacter sphaeroides*. *Micro. Biotechnol.* **13**, 1082–1093 (2020).
93. Orsi, E. et al. Growth-uncoupled isoprenoid synthesis in *Rhodobacter sphaeroides*. *Biotechnol. Biofuels* **13**, 123 (2020).
94. Lu, P. et al. Regulation of carbon flux and NADH/NAD⁺ supply to enhance 2,3-butanediol production in *Enterobacter aerogenes*. *J. Biotechnol.* **358**, 67–75 (2022).
95. Qiang, S. et al. Promoting the synthesis of precursor substances by overexpressing hexokinase (Hxk) and hydroxymethylglutaryl-CoA synthase (Erg13) to elevate β -carotene production in engineered *Yarrowia lipolytica*. *Front. Microbiol.* **11**, 1346 (2020).
96. Roberts, D. J. & Miyamoto, S. Hexokinase II integrates energy metabolism and cellular protection: acting on mitochondria and TORC1 to autophagy. *Cell Death Differ.* **22**, 248–257 (2015).
97. Robey, R. B. & Hay, N. Mitochondrial hexokinases, novel mediators of the antiapoptotic effects of growth factors and Akt. *Oncogene* **25**, 4683–4696 (2006).
98. Wilson, J. E. Hexokinases. In *Reviews of Physiology, Biochemistry and Pharmacology*, Vol. 126, 65–198 (Springer Berlin Heidelberg, 1995).
99. Um, J. et al. Metabolic engineering of *Enterobacter aerogenes* for 2,3-butanediol production from sugarcane bagasse hydrolysate. *Bioresour. Technol.* **245**, 1567–1574 (2017).
100. Kim, D. G. et al. Improved 2,3-butanediol yield and productivity from lignocellulose biomass hydrolysate in metabolically engineered *Enterobacter aerogenes*. *Bioresour. Technol.* **309**, 123386 (2020).
101. Rinaldi, M. A., Ferraz, C. A. & Scrutton, N. S. Alternative metabolic pathways and strategies to high-titre terpenoid production in *Escherichia coli*. *Nat. Prod. Rep.* **39**, 90–118 (2022).
102. Huang, K. X., Scott, A. I. & Bennett, G. N. Overexpression, purification, and characterization of the thermostable mevalonate kinase from *Methanococcus jannaschii*. *Protein Expr. Purif.* **17**, 33–40 (1999).
103. Primak, Y. A. et al. Characterization of a feedback-resistant mevalonate kinase from the archaeon *Methanosarcina mazei*. *Appl Environ. Microbiol.* **77**, 7772–7778 (2011).
104. Kazieva, E. et al. Characterization of feedback-resistant mevalonate kinases from the methanogenic archaeons *Methanosaeta concilii* and *Methanocella paludicola*. *Microbiology* **163**, 1283–1291 (2017).
105. Chen, H. et al. Enhancement of the catalytic activity of Isopentenyl diphosphate isomerase (IDI) from *Saccharomyces cerevisiae* through random and site-directed mutagenesis. *Micro. Cell Fact.* **17**, 65 (2018).
106. Shukal, S. et al. Metabolic engineering of *Escherichia coli* BL21 strain using simplified CRISPR–Cas9 and asymmetric homology arms recombineering. *Micro. Cell Fact.* **21**, 19 (2022).
107. Unger, T. et al. Applications of the restriction free (RF) cloning procedure for molecular manipulations and protein expression. *J. Struct. Biol.* **172**, 34–44 (2010).
108. Chen, X. et al. Total enzymatic synthesis of cis- α -iron from a simple carbon source. *Nat. Commun.* **13**, 7421 (2022).
109. Wada, K. et al. (¹³C)-metabolic flux analysis for mevalonate-producing strain of *Escherichia coli*. *J. Biosci. Bioeng.* **123**, 177–182 (2017).
110. Ng, P. et al. Optimization of extraction conditions for LC–ToF–MS analysis of mevalonate pathway metabolites in engineered *E. coli* strain via statistical experimental designs. *Talanta* **254**, 124182 (2023).
111. Nanchen, A., Fuhrer, T. & Sauer, U. Determination of metabolic flux ratios from ¹³C-experiments and gas chromatography-mass spectrometry data. In *Metabolomics: Methods and Protocols* (ed. Weckwerth, W.) 177–197 (Humana Press, 2007).
112. Chassagnole, C. et al. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol. Bioeng.* **79**, 53–73 (2002).

113. Millard, P., Smallbone, K. & Mendes, P. Metabolic regulation is sufficient for global and robust coordination of glucose uptake, catabolism, energy production and growth in *Escherichia coli*. *PLoS Comput. Biol.* **13**, e1005396 (2017).

Acknowledgements

This work was supported by the Intra-create Thematic Grant “Cities” (grant number: NRF2019-THE001-0007) under the EcoCTs project. The EcoCTs research project is supported by the National Research Foundation, Prime Minister’s Office, Singapore, under its campus for Research Excellence and Technological Enterprise (CREATE) programme. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. In addition, the authors are thankful to Dr. Florian Bellvert, Hanna Kulyk, and Cecilia Berges from MetaToul (Metabolomics & Fluxomics Facilities, Toulouse, France) for their experimental guidance and insights. We would also like to acknowledge Dr. Chew Wee, Dr. Derek Smith, and Shreyash Borkar from the Singapore Institute of Food and Biotechnology Innovation (SIFBI), A*STAR, Singapore for technical support and discussion. The software Biorender (Biorender.com) was utilised for some graphics (indicated in figure legends).

Author contributions

J.K.K. carried out culturing and metabolomics experiments on the EcoCTS03 and modified strains, as well as ¹³C-tracer studies. J.K.K. and Y.T.H. developed the dynamic model. C.P.M.S. and X.C. designed and prepared the modified strains. J.K.K. and K.S. designed the study and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-024-00440-7>.

Correspondence and requests for materials should be addressed to Kumar Selvarajoo.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024