Enhanced Production of (+)-Limonene through Targeted Engineering of *Citrus sinensis* Limonene Synthase

Clement P.M Scipion[1,2], Jérémy Esque[3], Shreyash Borkar[2], Ong Jing Sen[2], Cristalle Seah[1,2], Ong Leonard[2], Pnelope Ng[2], Yoganathan Kanagasundaram[2], Magali Remaud-Simeon[3], Sophie Bozonnet[3], Bo Xue[4,5,6] Wen Shan Yew[4.5.6], Isabelle André[3,], Xixian Chen[2,*]*

1. CNRS@CREATE, 1 Create Way, #08-01 Create Tower, Singapore 138602
2. Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 31 Biopolis Way, Nanos, Singapore 138669
3. Toulouse Biotechnology Institute, TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France. 135, avenue de Rangueil, F-31077 Toulouse Cedex 04, France
4. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore (NUS), 8 Medical Drive, Singapore 117597, Singapore
5. NUS Synthetic Biology for Clinical and Technological Innovation, 28 Medical Drive, Singapore 117456, Singapore

6. Synthetic Biology Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, 14 Medical Drive, Singapore 117599, Singapore

Corresponding Authors

* Corresponding authors: isabelle.andre@insa-toulouse.fr; **xixian_chen@sifbi.a-star.edu.sg**

Abstract

Limonene is a high-value monoterpene with numerous industrial applications, including flavors, fragrances, and pharmaceuticals. Its biosynthesis in microbes is often limited by the limonene synthase and can be enhanced through targeted enzyme engineering. This study aims to optimize the *Citrus sinensis* limonene synthase (CsLS) in order to boost (+)-limonene yield. Site-saturation mutagenesis was carried out to create a library of truncated CsLS variants, which were evaluated for enzymatic activity, enantioselectivity, and kinetic properties. Through medium-throughput screening, a mutant, tCsLS-Q8K, was found to improve limonene production by 2.1-fold. Bioreactor experiments with this optimized mutant, in combination with the *ispA* S80F mutation to enhance monoterpene precursor availability, achieved a (+)-limonene titer of 4.9 g/L, representing the highest production level of limonene reported to date. This work not only deepens our understanding of CsLS function but also offers a scalable approach for enhanced (+)-limonene production. The findings underscore the potential of enzyme engineering and pathway optimization to transform biotechnological processes, significantly impacting the commercial viability of the flavor and fragrance industries.

KEYWORDS: monoterpene, limonene, enzyme engineering, limonene synthase, *Citrus sinensis*

Introduction

Limonene, a naturally occurring monoterpene, is found in citrus fruits and in a plethora of plants[1], contributing to their scents[2]. It is widely used in the fragrance, food, cosmetics and pharmaceutical industries due to its pleasant citrus aroma and diverse functional properties such as anti-inflammatory, antioxidant, antiviral, neuroprotection and gastroprotective properties[3–5]. Additionally, it has emerging applications in the agri-food industry due to its herbicidal, antimicrobial and antioxidant properties[6], and could be used as a preservative[7–9]. The market for limonene is expected to grow due to its potential applications in renewable energy and bioplastics [10,11]. Limonene derivatives such as perillyl alcohol, caverol and linalool also display interests and application in these industries and offer novel applications such as jet fuels[12–15]. Limonene is currently mainly produced by extracting citrus peels, as a by-product of juice industries[1,16], which can be affected by diseases[17] and pesticides[18]. Sustainable and bio-based production of limonene using microbes becomes a promising alternative, which is pesticide-free, resilient to diseases and scalable to meet the increasing demand. In addition, bioproduction uses renewable substrates, ensures higher production specificity and operates with mild process conditions. Furthermore, bioproduced limonene can be considered natural[1]. Therefore, enhancing the efficiency of limonene production not only meets current industrial demands but also supports the development of sustainable and eco-friendly production processes.

The biosynthesis of limonene in plants occurs through the mevalonate (MVA) or the methylerythritol phosphate (MEP) pathway. The MVA pathway has been previously introduced in *Escherichia coli* for the production of limonene[19–21]. In this pathway, acetyl-CoA undergoes a series of enzymatic reactions to form mevalonate, which is then converted into isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), which are terpene precursors

(**Figure 1**). These are further condensed to generate geranyl pyrophosphate (GPP). The *cis* isomer of GPP, neryl pyrophosphate (NPP), has been reported to be a natural alternative for the production of terpenes[22]. An NPP synthase from *S. lycopersicum* (slNPPS) was described to produce NPP from IPP and DMAPP[22]. NPP is not metabolized in heterologous microbial hosts such as *E. coli,* unlike GPP, which can be converted to farnesyl pyrophosphate (FPP) by FPP synthase. This leads to the accumulation of NPP in microbial hosts, which could be beneficial for the yield improvement. GPP or NPP is further converted to limonene by limonene synthases (LSs). Several enantiospecific LSs have been reported[23–27], such as *Mentha spicata* LS which is used for the production of (-)-limonene[12,20], and *Citrus sinensis* or *Citrus limon* for the production of (+)-limonene[28,29]. LSs are type I cyclases[30], and the catalytic N-terminal domain of LS contains an active site with a DDxxD motif that is highly conserved across monoterpene synthases[31]. The aspartate residues bind to a trinuclear cluster of divalent metal ions (Mg^{2+} or Mn^{2+}). This cluster helps to form a complex with the leaving pyrosphosphate moiety from GPP or NPP, resulting in the creation of characteristic carbocation intermediates[32]. The catalytic domain also contains a C-terminal RR(X_8)W motif.

LSs have been overexpressed in *E. coli* and *Saccharomyces cerevisiae* for limonene production from GPP or NPP[19,33] and are often considered to be the rate-limiting step due to its suboptimal catalytic efficiency[34,35]. Previous studies have shown that the removal of the N-terminal sequence of LS could significantly increase its activity[36]. Specially, truncating the LS before the arginine tandem R58-R59 (*Mentha spicata* LS) improved the limonene production, while truncation after R58 drastically reduced the limonene production[36]. This suggests that engineering of LS could be a promising way to enhance limonene production yield. Several studies have explored various mutants of limonene synthases to enhance production efficiency. A study utilizing the orthogonal

NPP pathway led to enhanced monoterpene production by engineering *Citrus limon* LS (H570Y), achieving 4.7-fold improvement compared to wild-type enzyme with the GPP pathway[33]. Other mutants, H570V, H570I and H570L lowered K_m value and increased efficiency for NPP compared to GPP[33,37]. Another work revealed that mutations in conserved regions of *Mentha spicata* limonene synthase, G566A and L571F, improved limonene production by 1.7-fold[38]. Using site-directed mutagenesis to investigate substrate specificity in *Citrus sinensis* limonene synthase, Schiff and colleagues showed that the majority of mutations targeting the active site caused a loss of activity with both GPP and NPP[39], except H571F and H571Y gave rise to only marginal improvements with GPP and NPP, respectively[39]. Another study explored the structure−function relationships in limonene synthases by identifying key active site residues that influence enantioselectivity. By mutating *Mentha spicata* (-)-limonene synthase (C321S, N345I, I453V, and M458V), the enantioselectivity was reversed to produce predominantly (+)-limonene (80%), enhancing understanding of monoterpene enantiomer formation[32].

Despite engineering efforts, the highest titers of limonene were achieved using wild-type limonene synthases. Willrodt and colleagues optimized the synthesis of (-)-limonene using recombinant *E. coli* with glycerol and glucose as carbon sources, employing limonene synthase from *Mentha spicata*[40]. They achieved a maximum (-)-limonene titer of 1.35 g/L in a two-liquid phase fed-batch fermentation, with glycerol improving both growth and production phases compared to glucose. Rolf and colleagues used *E. coli* BL21 (DE3), metabolically engineered with pJBEI-6410[20] plasmid, which contained optimized mevalonate pathway genes and *Mentha spicata* limonene synthase, in a fed-batch bioreactor setup, utilizing glycerol as the sole carbon source[41]. They achieved a limonene titer of 3.6 g/L and a space-time yield of 0.15 g/L/h, representing the highest limonene concentrations reported to our knowledge using microorganisms.

In this research, we aimed to improve the production of (+)-limonene by engineering *Citrus sinensis* LS (CsLS). Truncated CsLS (tCsLS) displays a 305-fold (+)-limonene production improvement compared to CsLS (**Figure S1**) **.** Our approach involved constructing a library of tCsLS mutants through site-saturation mutagenesis and assessing them using a medium-throughput system. The most promising mutants were then further tested and optimized under controlled conditions in flasks and subsequently scaled up in a bioreactor. This methodology led to a significant improvement in limonene titer, reaching 4.9 g/L, the highest level for (+)-limonene reported to date. This study demonstrates the potential of enzyme engineering to overcome the rate-limiting step in limonene biosynthesis, thereby achieving unprecedented production levels. Our findings contribute to the broader effort of developing efficient biotechnological solutions for the sustainable production of high-value compounds.

Results and discussion

tCsLS site-saturation mutagenesis

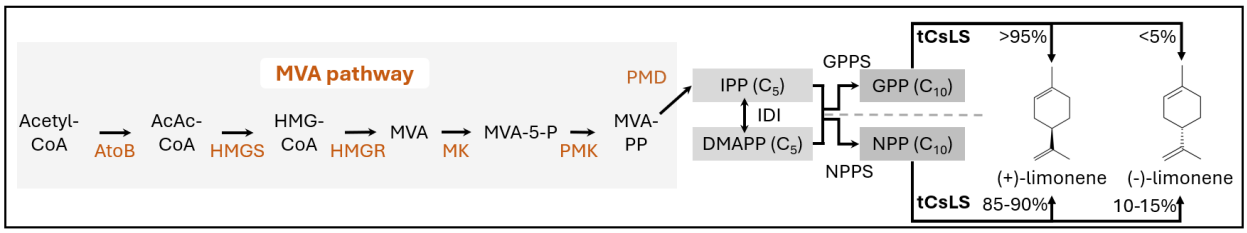

Figure 1. Biosynthetic pathway of limonene using *Citrus sinensis* limonene synthase (tCsLS). Terpene precursors isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are synthesized through the mevalonate (MVA) pathway from acetyl-CoA. These are converted to geranyl pyrophosphate (GPP) by the GPP synthase (GPPS) enzyme, or to neryl pyrophosphate (NPP) by NPP synthase (NPPS). GPP or NPP are further converted to limonene by *Citrus sinensis* limonene synthase (tCsLS). AtoB, Acetyl-CoA acetyltransferase; AcAc-CoA, acetoaceyl-CoA; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMG-CoA, 3-hydroxy-3-methylglutaryl-

CoA; HMGR, HMG-CoA reductase; MK, mevalonate kinase; MVA-5-P, mevalonate-5-phosphate; PMK, phosphomevalonate kinase; MVA-PP, mevalonate-5-diphosphate; PMD, mevalonate diphosphate decarboxylase; IDI, isopentenyl-diphosphate isomerase.

To screen truncated *Citrus sinensis* limonene synthase (tCsLS) mutants, we designed and optimized a protocol to automate cloning, transforming, cell culturing, and analysis. We selected 12 amino acid residues for site-saturation mutagenesis from the following three analyses (**Figure 2a**): (i) residues structurally located near the active site (S261, I285, I394, T395, L398, I399, I400, F433); (ii) residues suggested by the HotSpot Wizard 3.1 Web Server[42] for potential stabilizing mutations (G449 and T527); (iii) residues found by sequence analysis (Q8 and Q525). For this latter selection, sequence alignment was performed from several limonene synthases, notably *Agastache rugosa* and *Nepeta tenuifolia* limonene synthases (tAgLS and tNtLS) which share similar active site composition but differ in substrate specificity and limonene production (data not shown). Of note, the sequence numbering corresponds to the truncated form for each enzyme.

Using our in-house integrated biofoundry SPARROW (SIFBI PARallel RObotic Workstation), we performed site-saturation mutagenesis at 12 positions and obtained 228 single mutants in the pEcoCTs-CPMS5 plasmid (**Table S1**). These plasmids were co-transformed with either pEcoCTs-CPMS6a (GPP pathway, **Table S1**) or pEcoCTs-CPMS6b (NPP pathway, **Table S1**) and cultured in 96 deep well plates with 20% isopropyl myristate (IM) for three days. Cultures were then measured for OD, and the IM layer was extracted, diluted in isopropanol, and analyzed by LC-MS. Each mutant was cultured in duplicate, with wild-type tCsLS as controls included on each plate. The OD-specific limonene titer of each mutant was normalized to the wild-type titer on the same plate. As shown in **Figure 2b**, most mutants exhibited low limonene production, especially

those corresponding to mutations near the active site, apart from I399 and I400. Interestingly, mutants based on sequence analysis (Q8 and Q525) revealed several mutations leading to improved titers compared to wild-type enzyme.

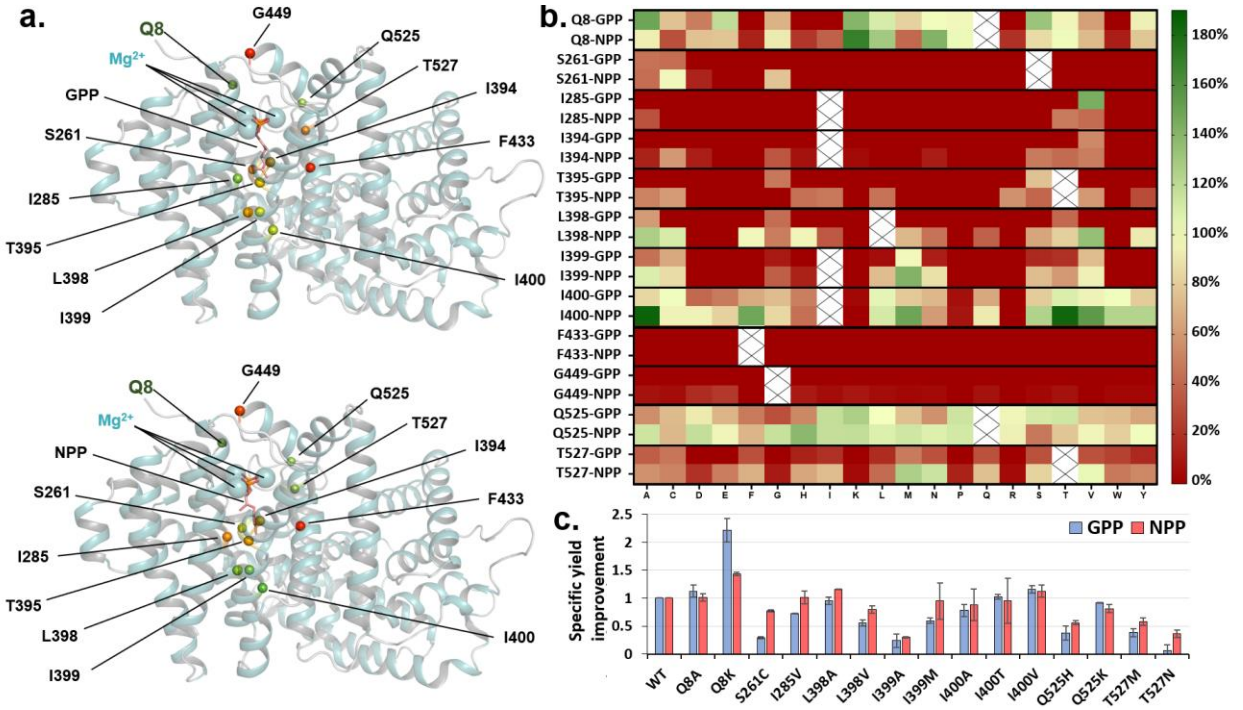

Figure 2. Improvements in (+)-limonene production through Site-Saturation Mutagenesis of tCsLS. (a) Mapping of site-saturated positions on the 3D structure of tCsLS in complex with GPP (on the top) and NPP (on the bottom). The enzyme is shown as ribbons, whereas GPP/NPP are displayed as sticks. The saturated positions are highlighted by spheres colored using the same color code as in (b). (b) Heatmap showing the yield improvement of tCsLS mutants compared to wild-type (WT) after site-saturation mutagenesis (SSM) on 12 selected sites. Culture was performed in deep well plates and analyzed by LC-MS (c) Limonene specific yield improvements of the best mutants identified from the SSM screening in (b), grown in tubes and analyzed by GC-MS.

To confirm these results, the top 15 mutants with improved limonene production were further cultivated in R-media with a dodecane second-phase layer 14 mL in snap-cap tubes. The strain used were MG1655-CPMS-6a5 or MG1655-CPMS-6b5 for GPP and NPP pathway, respectively (**Table S2**). The dodecane layer was then extracted and diluted in hexane, followed by analysis using GC-MS. Here, we confirmed that the tCsLS-Q8K mutant exhibited a 2.1-fold increase in limonene production via the GPP pathway and a 1.4-fold increase via the NPP pathway (**Figure 2c**). Unfortunately, most of the other mutants (93%) were false positives, probably due to the high standard deviations obtained during the medium-throughput screening (**Figure S2a**). From these results, several double mutants were generated by recombining mutations, but they showed no further improvement in limonene production compared to the Q8K mutant (**Figure S2b**). It is worth noting that variability between the deep well plate (**Figure S2a**) and 14 mL tube (**Figure 2c**) experiments, analyzed by LC-MS and GC-MS respectively, can lead to discrepancies, as observed with the Q525K mutant, which showed higher limonene production in the screening but a lower specific yield in the subsequent tests.

To check for the enantiospecificity of the engineered LSs, (+) and (-)-limonene produced by wild-type tCsLS and the 15 mutants were analyzed using a chiral column (**Figure S3**). Notably, it was found that wild-type tCsLS produced (+)-limonene with a 96% enantiopurity through the GPP pathway and only 4% of (-)-limonene. Interestingly, the same enzyme produced higher amounts of (-)-limonene (12%) via the NPP pathway, a trend also observed for the 15 other mutants yielding between 11% and 21% of (-)-limonene. In addition, the I399A and T527N mutations were found to modify the enantioselectivity, as they produce 20% and 27% (-)-limonene through GPP pathway, respectively. This result suggests the potential to engineer a switch for (-)-limonene production using a (+)-limonene LS, similar to the previously engineered (-)-LS that produces (+)-

limonene[32]. Importantly, the tCsLS-Q8K mutant maintained its enantioselectivity via the GPP pathway, producing only 3% (-)-limonene.

Characterization of tCsLS-Q8K

To further investigate the properties of tCsLS-Q8K, we purified both the wild-type tCsLS and the Q8K mutant and conducted a kinetic analysis. The kinetic data revealed trends consistent with our results *in vivo.* For both substrates (GPP and NPP), a decrease in K_m and an increase in k_{cat} were observed for the Q8K mutant compared to the wild-type enzyme, which resulted in an increase in the catalytic efficiency of the mutant tCsLS-Q8K by a factor of 2.2 and 2.65 for the GPP and NPP substrates respectively, compared *to* tCsLS. (**Figure 3a and Figure S4**).

Moreover, The Q8 residue is located at the surface of the enzyme, near the active site, and belonging to the RR(X_8)W motif[43]. Introduction of a positive charge (K) at protein surface and near N-terminal has been shown to improve the protein expression in some studies[44–48]. We then tested the effect of Q8K mutation on expression by transforming *E. coli* using only the pEcoCTs-CPMS5 plasmid, carrying either the wild-type LS or Q8K mutant. The cultures were grown using the same protocol used for mutant characterization described in the previous section, *i.e.* for three days, in snap-cap tubes, with a dodecane layer. SDS-PAGE was performed on the cell lysates, and the results showed a clear increase in the expression of the Q8K mutant compared to the wild-type (**Figure 3b and Figure S5**). Moreover, this increased expression level could likely enhance the overall enzyme activity, leading to greater substrate turnover and higher product yield. Quantification of purified enzyme showed a ~10% higher yield for Q8K compared to the wild type, further supporting improved expression. These results suggest that higher enzyme expression contributes to the improved performance of the Q8K mutant. Increased enzyme levels can drive

higher substrate consumption and more efficient limonene biosynthesis, ultimately leading to higher product yields. These findings corroborate our initial observations and provide a deeper understanding of the reasons underlying the improved functionality of the tCsLS-Q8K mutant.

a.

	tCsLS		tCsLS-Q8K	
	GPP	**NPP**	**GPP**	**NPP**
K_m (μM)	46.5 ± 0.3	24.7 ± 2.0	34.5 ± 0.8	22.5 ± 1.8
k_{cat} (sec^{-1})	0.019 ± 0.000	0.012 ± 0.002	0.031 ± 0.000	0.029 ± 0.002
Catalytic efficiency K_m/K_{cat} (mM^{-1} sec^{-1})	0.41 ± 0.01	0.49 ± 0.04	0.90 ± 0.03	1.3 ± 0.01

b.

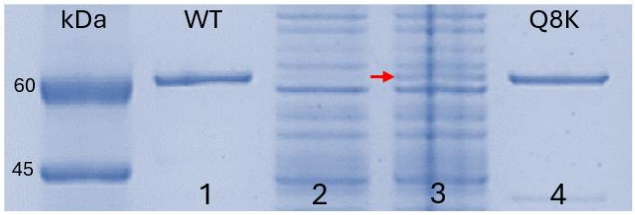

Figure 3. Kinetics and expression analysis of wild-type tCsLS (WT) and mutant tCsLS-Q8K. (a) Kinetic parameters of wild-type tCsLS and the tCsLS-Q8K mutant. (b) SDS-PAGE analysis of purified tCsLS (lane 1, WT) and tCsLS-Q8K (lane 4, Q8K), and the expression levels of both WT (lane 2) and Q8K mutant (lane 3) in total cell extracts. The red arrow indicates the expression band for tCsLS.

Having established the promising potential of the Q8K mutant, our next step involved scaling up the production to flask cultures and optimizing the conditions to further enhance limonene yield. This transition to larger-scale production will allow us to evaluate the practical applications of the Q8K mutant and identify any additional factors that could influence its performance.

Flask optimization to improve limonene production

To maximize (+)-limonene production, we developed a two-plasmid system and performed CRISPR modifications on *E. coli* MG1655 strain, and GPP pathway was used, since it produces predominantly (+)-limonene. This section details our systematic optimization of various parameters in shake flask cultures, including shaking speed, carbon source selection, and the concentration of the inducer. By fine-tuning these conditions, we aimed to enhance the overall efficiency and yield of (+)-limonene production in preparation for scaling up the process to a bioreactor.

We first developed a two-plasmid system to optimize the biosynthesis of (+)-limonene, as illustrated in **Figure S6**. The first plasmid, pEcoCTs-CPMSN1, contains the genes encoding the enzymes of the mevalonate pathway, ensuring an efficient supply of precursors. The second plasmid, pEcoCTs-CPMS2, harbours the genes for geranyl pyrophosphate synthase (GPPS) and tCsLS-Q8K, which are converting these precursors into (+)-limonene.

To further enhance production, we performed CRISPR-Cas9 recombineering[49] on the *E. coli* strain MG1655 genome to introduce the S80F mutation in the *ispA* gene. This mutation in *ispA* was previously reported to redirect the metabolic flux towards monoterpenes in *E. coli*, allowing GPP to accumulate[12,50]. We compared the wild-type strain with the *ispA* S80F mutant strain in shake flask cultures and found 79% improvement in the limonene titer with the mutant strain, as shown in **Figure S7a**. Given the significant improvement in limonene titer, the *ispA* S80F strain was used for all subsequent flask cultures (Strain MG1655ispA-N12Q8K, **Table S2**).

Using this strain, we optimized the shaking speed to enhance (+)-limonene production. During this phase, we employed lactose induction, which was also used during the site-saturation mutagenesis initial screening. Previously, we have seen that monoterpene production is sensitive

to oxygen level[51], thus we compared limonene production at two shaking speeds: 100 rpm and 300 rpm. The results showed that 300 rpm yielded a significant 2.6-fold increase in limonene production compared to 100 rpm (**Figure 4a**), highlighting that an improved oxygen supply with higher shaking speed could be beneficial for limonene production. Moreover, we supplemented the cultures with an amino acid cocktail, but this did not result in additional improvement (**Figure 4a**). Consequently, we concluded that the higher shaking speed, providing better oxygenation, was optimal for limonene production.

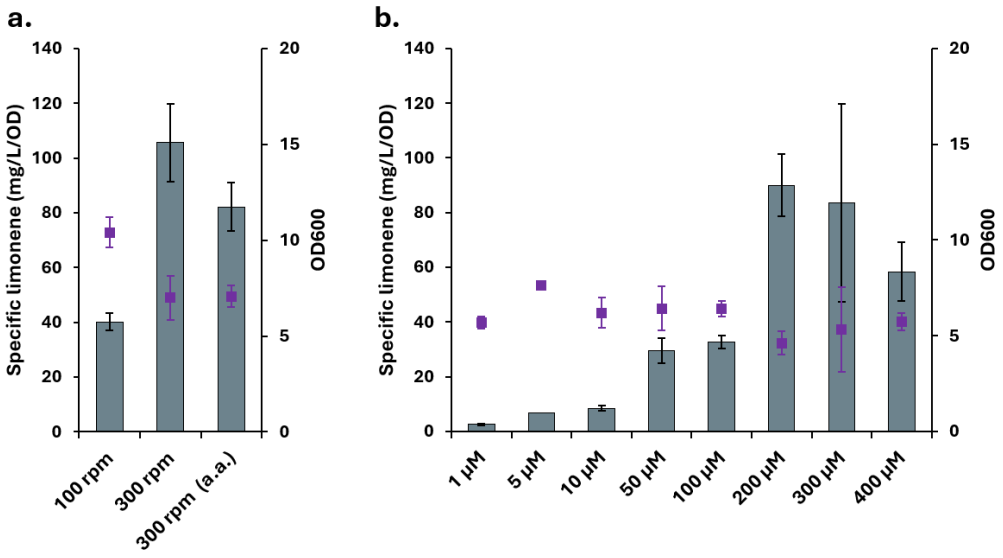

Figure 4. Optimization of shake flask conditions for enhanced (+)-limonene production. (a) Effect of shake speed and amino acid (a.a.) supplementation on (+)-limonene production. (b) Optimization of IPTG concentrations for inducing the pathway gene expression. OD_{600} is represented in purple.

Next, IPTG induction was tested, as autoinduction media is not easily scalable. We optimized IPTG concentration while maintaining optimal conditions, namely shake flasks with the MG1655ispA-N12Q8K strain, 300 rpm shaking speed, and using glycerol as the carbon source.

Since previous studies identified glycerol as a superior carbon source for monoterpenes and limonene production[40,41,52], we first tested its performance in our system. Under 200 μM IPTG induction, indeed, a higher specific titer for limonene was obtained using glycerol substrate, although with no statistical difference compared to that with glucose (**Figure S7b**). We then varied IPTG concentrations from 1 to 400 μM. We found that IPTG at 200 μM not only produced the highest limonene titer but also the most reproducible results (**Figure 4b**). Notably, the limonene titers obtained with 200 μM IPTG were comparable to those obtained with autoinduction media, indicating that IPTG induction at this concentration is effective for scaling up production.

Building on these optimization steps, we next compared WT tCsLS against the Q8K tCsLS mutant under the best-performing shake flask conditions (*ispA* S80F strain, 300 rpm, 10 g/L glycerol feed, and 200 μM IPTG). The Q8K mutant produced significantly more limonene than the WT, confirming that the titer increase is primarily driven by the mutation rather than the conditions alone (**Figure S7c**), achieving a 2-fold improvement. These optimal conditions were employed in the bioreactor stage of our process.

Limonene production in a fed-batch process

Building on the optimized conditions established in shake flasks, we scaled up the process to a 5 L bioreactor using the MG1655ispA-N12Q8K strain. The initial media consisted of R-media with 20 g/L glycerol and a volume of 1.8 L, supplemented with 68 μg/mL chloramphenicol and 200 μg/mL ampicillin. After reaching an optical density (OD) of 10, induction was carried out with 200 μM IPTG. To facilitate the extraction of limonene, 1 L of dodecane was added, and glycerol feeding commenced. The feed solution comprised 800 g/L glycerol and 8 g/L $MgSO_4$, with a feeding rate of 12.7 mL/h. Following induction, the temperature was adjusted to 30°C, while the dissolved oxygen (DO) level was maintained at 30% and the pH at 7.

The bioreactor run was conducted for 70 hours, during which the maximum limonene titer reached 4.9 g/L, with a peak yield of 3.2% achieved at 65 hours (**Figure 5**). The production rate was 114 mg/L/h. The enzyme remained active throughout production, as limonene levels continued to increase up to ~65 hours, indicating sustained activity during the culture. The optimized conditions, including specific IPTG induction, glycerol feeding, and dodecane addition, were crucial in maximizing yield and production rate. The observed accumulation of glycerol and acetate indicates that there is potential for further strain and process improvement. Nonetheless, this represents the highest titer achieved to date, demonstrating the effectiveness of both our engineered limonene synthase and the optimized conditions in the bioreactor setting, and its potential for industrial-scale production.

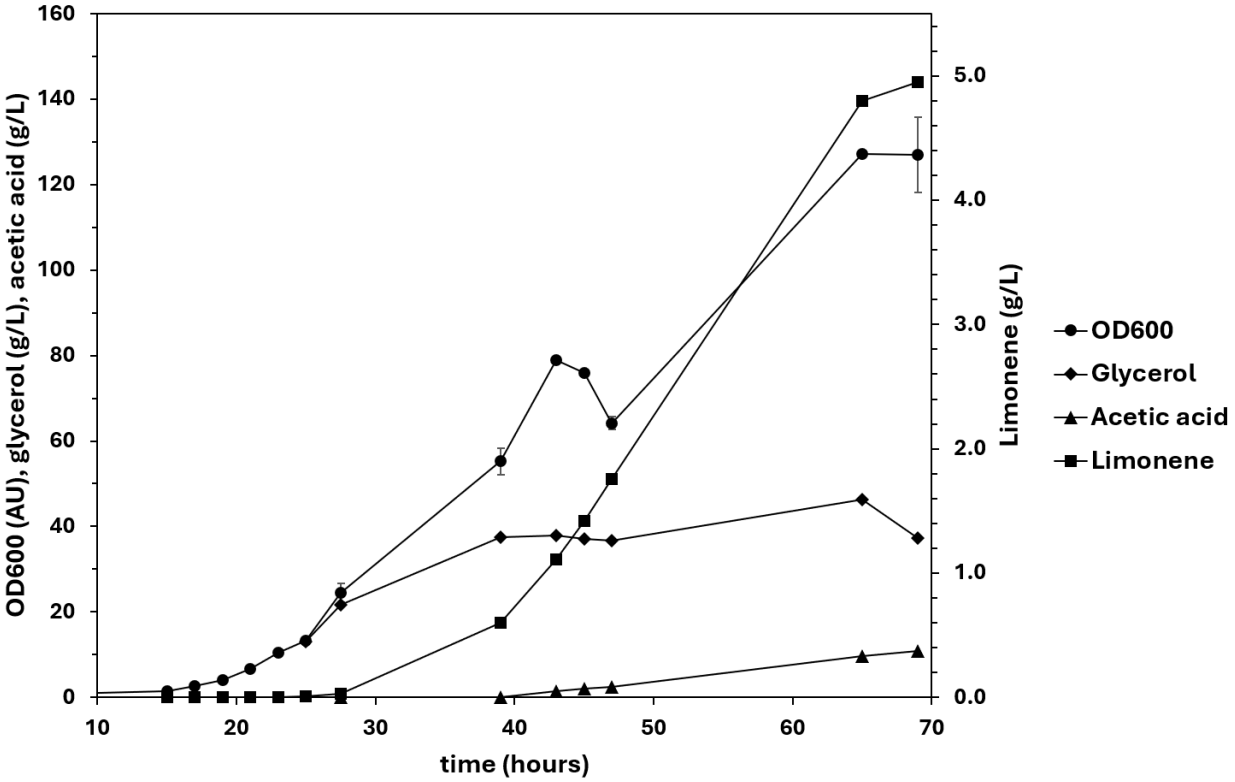

Figure 5. Fed-batch production of (+)-limonene in a 5 L bioreactor using optimized tCsLS-Q8K.

4. Conclusions

This study successfully demonstrates the potential of enzyme engineering combined to process optimization to enhance (+)-limonene production. The site-saturation mutagenesis of tCsLS allowed us to identify the tCsLS-Q8K mutant which significantly enhanced limonene production and maintained enantioselectivity. Detailed kinetic analysis confirmed that the Q8K mutant exhibited increased catalytic efficiency for GPP and NPP. Furthermore, SDS-PAGE analysis demonstrated higher expression levels of the Q8K mutant compared to the wild-type enzyme, which likely contributed to its better performance. We then optimized several parameters for (+)-limonene production in shake flasks. The two-plasmid system, containing tCsLS-Q8K, was used in conjunction with the *ispA* S80F strain. Optimal conditions included a shaking speed of 300 rpm, glycerol as the carbon source, and an IPTG concentration of 200 μM. These conditions have significantly enhanced limonene production and were carried forward to the bioreactor stage, where a titer of 4.9 g/L was achieved.

These advancements mark a significant improvement over previous production methods and highlight the feasibility of engineering limonene synthase for industrial-scale applications. The results underscore the importance of integrating enzyme engineering with optimized fermentation processes to achieve high production yields. Indeed, the combination of the tCsLS-Q8K mutant with the *ispA* S80F mutation contributed to the significant improvement in limonene production. Overall, this study lays a strong foundation for advancing the sustainable production of (+)-limonene and contributes valuable insights into the broader field of enzyme engineering and

bioprocess optimization. Furthermore, optimizing downstream process, using hydrodistillation or chromatography, for example, will be essential to improve product recovery and large-scale limonene production. Continued research in these areas will pave the way for further innovations and improvements in the production of bio-based chemicals.

Experimental procedures

Plasmids, strains and chemicals

The plasmids pEcoCTs-CPMS6a, 6b and 5, which were used for medium-throughput screening, were prepared as described in [**our manuscript under review**]. **pEcoCTs-CPMS6a** was created by deleting the tcr promoter, trGPPS, LS and rrnB T1 and T2 terminators from pJBEI-6409[20] plasmid using restriction-free cloning[53]. This intermediate plasmid is pEcoCTs-CPMSN1 (**Figure S6**) and contains the enzymes of the mevalonate pathway. pEcoCTs-CPMSN1 was used as a backbone for the insertion of the tcr promoter, trGPPS, LS and rrnB T1 and T2 terminators in-between the p15A origin of replication and the chloramphenicol acetyltransferase expression site. Finally, LS was deleted from the second intermediate plasmid using restriction-free cloning, to obtain pEcoCTs-CPMS6a. pEcoCTs-CPMS6a contains enzymes of the mevalonate pathway and GPPS (**Figure 1**). pEcoCTs-CPMS6b was created by using pEcoCTs-CPMS6a without GPPS as a backbone and truncated NPPS from *S. lycopersicum* as an insert, and 2-piece cloning was performed. pEcoCTs-CPMS6b contains enzymes of the mevalonate pathway and NPPS. pEcoCTs-CPMS5 was created by deleting GPPS from the pJBEI-3101 plasmid. Full-length *Citrus sinensis* limonene synthase (CsLS) was obtained from Bio Basic cloned in pCPMS-EcoCTs5. Truncated CsLS (tCsLS) in pEcoCTs-CPMS5 was obtained by restriction free cloning, by designing primers with overhangs that contained the desired deletion in the overlap region. The

pEcoCTs-CPMS2 plasmid was obtained by substituting *Mentha spicata* limonene synthase by truncated *Citrus sinensis limonene* synthase (tCsLS) on the pJBEI-3101[20] plasmid.

For protein overexpression and purification, the genes of interest were cloned in psY7, a vector derived from pET-21d (Novagen). It incorporates an N-terminal maltose binding protein (MBP) tag, a His-tag, followed by an HRV 3C protease site and the protein of interest. Plasmids psY7-tCsLS and -tCsLS-Q8K were obtained by 2-piece cloning, using psY7 without the protein of interest as a backbone, and tCsLS and tCsLS-Q8K as inserts.

The list of plasmids and strains are provided in **Tables S1 and S2**.

Preparation of MG1655 *ispA* S80F by CRISPR

To perform the mutation of *ispA* genes S80F from K12 *E. coli* MG1655 strain, CRISPR Cas9-assisted recombineering method was used[49]. gRNA sequences were identified for wild-type Cas9 of 20bp length in the target gene with reference to the MG1655 genome - ASM584v2 (*E. coli* str. K-12 substr. MG1655) for SpCas9 protospacer adjacent motif (PAM), NGG on 3' side. A guide RNA sequence covering the site of mutagenesis was selected. p15A origin pTarget plasmid used in our previous work was used[49]. Plasmid pTarget-J23119- ispA-MG1655-HA-ispA-S80F was assembled using In-Fusion plasmid DNA assembly with pTarget plasmid, gRNA and homologous arm harboring mutation. CRISPR Cas9-assisted recombineering was then performed as previously described[49]. The strains were stored in 40% glycerol at -80°C before further use.

Site-saturation mutagenesis

228 tCsLS mutants were cloned from pEcoCTs-CPMS5 using restriction-free cloning[,53]. Forward primers containing the desired mutation and a 5’ 15-bp overlap with a single reverse primer were designed for each of the 12 selected SSM positions. A total of 240 primers were designed (19 forward and 1 reverse for each mutation site, **Table S3**). For each position, the Y

mutant was cloned manually for forward primer optimization, as the primer length was kept to a minimum to reduce cost. Subsequently, SPARROW automated biofoundry was used to perform the cloning of the remaining 18 mutants for each site (216 mutants). Each mutant was confirmed using Sanger sequencing. The 228 plasmids were stored in 96-well PCR plates at -20°C. The SPARROW system is an integrated robotic automation platform within our biofoundry designed for high- and medium-throughput screening of enzyme variants. It combines a range of automated instruments for colony picking, liquid handling, incubation, PCR, centrifugation, and analysis. Key components include the QPix420 colony picking system (Molecular Devices), the Biomek i7 liquid handler, and the PF-400 robot arm (Brooks) for efficient sample manipulation. The platform is designed to streamline the entire process from library generation to high-throughput screening, significantly increasing throughput and reducing manual effort.

Limonene production and analysis

General protocol. For limonene production experiments, competent cells of the *E. coli* K-12 MG1655 or MG1655 *ispA S80F* were prepared using the Mix & Go! *E. coli* Transformation Kit (Zymo Chem) following manufacturer's specifications. Plasmids were then co-transformed by heat shock and plated on a LB agar plate supplemented with chloramphenicol (34 µg/mL) and ampicillin (100 µg/mL) and incubated overnight at 37°C. Freshly grown colonies were picked and inoculated in LB media with chloramphenicol (34 µg/mL) and ampicillin (100 µg/mL), and incubated at 37°C, 300 rpm overnight. Each pre-culture was inoculated in R-media (2 g/L ammonium sulphate, 4.2 g/L KH_2PO_4, 11.24 g/L K_2HPO_4, 1.86 g/L citric acid, and 10 mL/L of 100x trace element solution. The 100x trace element solution consists of $CoCl_2 \cdot 6H_2O$ (0.25 g/L), $MnSO_4 \cdot 4H_2O$ (1.5 g/L), $CuSO_4 \cdot 2H_2O$ (0.15 g/L), H_3BO_3 (0.3 g/L), $Na_2MoO_4 \cdot 2H_2O$ (0.25 g/L), $Zn(CH_3COO)_2$ (0.8 g/L), Fe(III) citrate (5 g/L), and EDTA (0.84 g/L) at pH 8). Magnesium

sulphate solution containing magnesium sulphate (0.5 g/L) and thiamine (0.0045g/L) was supplemented to the media. The carbon sources were added, an inducer and two times the usual working concentration of chloramphenicol (68 µg/mL) and ampicillin (200 µg/mL) to a starting OD_{600} of 0.1. A second phase overlay was added to the media for limonene extraction. The media was then incubated at 30°C, 300 rpm for 72 hrs. After 72 hrs, the OD of the solution was measured, and the second-phase layer was extracted for further analysis. Limonene concentration was determined by GC-MS in mg/L. This concentration was divided by OD600 to determine the specific limonene concentration (mg/L/OD). Specific limonene concentration was divided by the WT control's specific limonene concentration to determine the specific yield improvement. (+) or (-) limonene (Sigma-Aldrich) were serial-diluted and used as standards for calibration. Details for each experiment are given below.

Site saturation mutagenesis (SSM). For SSM, using the automation system, plasmids pEcoCTs-CPMS6a or pEcoCTs-CPMS6b were co transformed in parallel with pEcoCTs-CPMS5. Two colonies were picked and incubated overnight in 500 µL LB in 96 deep well plates, then re-inoculated (OD 0.1) in 800 µL R-media in 96 deep well plates, using 10 g/L glycerol as a carbon source and 1 g/L glucose and 2.5 mM lactose as an auto-induction system. 160 µL (20%) isopropyl myristate (IM) was added to each well. After 72 hours, 20 µL R-media was extracted to measure OD_{600}, and 140 µL IM was added to each well, and 50 µL of the IM layer was extracted and diluted in isopropanol, followed by LC-QqQ analysis.

Analysis of Mutants. The best selected mutants were transformed using pEcoCTs-CPMS6a or pEcoCTs-CPMS6b and pEcoCTs-CPMS5. For total limonene production analysis, three colonies were picked and incubated overnight in 1 mL LB in 14 ml snap-cap tubes, then re-inoculated (OD 0.1) in 1 mL R-media in 14 mL snap-cap tubes, using 10 g/L glycerol as a carbon source and 1 g/L

glucose and 2.5 mM lactose as an auto-induction system. 200 µL (20%) dodecane was added to each tube. After 72 hours, OD_{600} was measured, and 50 µL dodecane was diluted in 950 µL hexane, followed by GC-MS DB-5 column analysis.

To analyze the enantiomers, one colony was picked and incubated overnight in 1 mL LB in 14 mL snap-cap tubes, then re-inoculated (OD 0.1) in 1 mL R-media in solid phase microextraction (SPME) vials, using 10 g/L glycerol as a carbon source and 1 g/L glucose and 2.5 mM lactose as an auto-induction system. No second-phase layer was used. After 72 hours, SPME GC-MS CycloSil-B chiral column analysis was performed.

Expression analysis. For the expression analysis of limonene synthases in the lysate, pEcoCTs-CPMS5 only was transformed. Two colonies were picked and incubated overnight in 1 mL LB in 14 ml snap-cap tubes, then re-inoculated (OD 0.1) in 1 mL R-media in 14 mL snap-cap tubes, using 10 g/L glycerol as a carbon source and 1 g/L glucose and 2.5 mM lactose as an auto-induction system. 200 µL (20%) dodecane was added to each tube. The media were only supplemented with ampicillin. After 72 hours, the cultures were spun down, the cell pellet was resuspended in 50 µL B-PER (ThermoFisher), vortexed for 10 minutes at room temperature, and 13 µL of the lysate was loaded into an SDS PAGE, alongside purified tCsLS and tCsLS-Q8K.

Shake flasks optimization. tCsLS or tCsLS-Q8K were transformed using pEcoCTs-CPMSN1 and pEcoCTs-CPMS2. Three colonies were picked and incubated overnight in 1 mL LB in 14 ml snap-cap tubes, then re-inoculated (OD 0.1) in 10 mL R-media in 250 mL shake-flasks using either autoinduction (10 g/L glycerol, 1 g/L glucose and 2.5 mM lactose) or IPTG induction (1 to 400 µM) with either glucose or glycerol 10 g/L. 10 mL (100%) dodecane was added to each flask. If necessary, Yeast Synthetic Drop-out Medium Supplements (Sigma Aldrich), 1.92 g/L, was added

to each flask. After 72 hours, OD_{600} was measured, and 50 µL dodecane was diluted in 950 µL hexane, followed by GC-MS DB-5 column analysis.

Fed-batch bioprocess. tCsLS-Q8K was transformed using pEcoCTs-CPMSN1 and pEcoCTs-CPMS2. Ten colonies were picked and incubated overnight in 1 mL LB in 14 ml snap-cap tubes. One colony was selected and re-inoculated in 250 mL LB in a 2 L shake-flask, and incubated overnight at 300 rpm, 37°C. Overnight culture was inoculated into 1.8 L of defined R-media with 20 g/L glycerol. Fermentation was conducted at 37°C and setpoint of dissolved oxygen (DO) was maintained at 30%. pH was controlled at 7 with 28% ammonium hydroxide. Cells were induced with 200 µM IPTG when OD_{600} reached 10. Following induction, temperature was lowered to 30°C and 1 L of dodecane was supplemented into the broth as in-situ extractant for limonene. Additional substrate dosed from a feed solution contains 800 g/L glycerol and 8 g/L $MgSO_4$. A continuous rate of 12.7 mL/h was fed and maintained until the end of the run at 69 hours. For each time point, 5 mL was extracted from the reactor, OD_{600} was measured, and dodecane was diluted in hexane, followed by GC-MS DB-5 column analysis.

LC-MS analysis of limonene. Chromatographic analysis was performed on an Agilent UPLC 1290 system coupled with an Agilent 6495C QqQ-MS with Agilent Jet Stream (AJS) Atmospheric Pressure Chemical Ionization (APCI) source operating in Multi Reaction Monitoring (MRM) mode. Separation of limonene was performed on an Agilent Poroshell C18 (50 mm x 4.6 mm x 5 µm), equipped with a 5mm guard column. The method utilized a dual mobile phase system: 0.1% formic acid in A) water and B) methanol at a constant flow rate of 0.5 mL/min. The elution gradient was as follows: 30% B at 0 min, 60% B at 1 min, 100% B at 4-5 mins and the column was re-equilibrated back to starting conditions in the next 2 mins. Column temperature was 30°C and injection volume was 2 µL. The APCI source parameters were as follows: gas temperature and

vaporizer: 290 °C; gas flow: 12 L/min, nebulizer pressure: 60 psig; capillary voltage: 2000 V; corona voltage: 10V; fragmentor voltage: 166V. Detection was performed in positive mode. Eluent was diverted to the MS from 1-5 min. The ion pairs that were monitored for limonene were 137.1-81.1 (quantitative ion) and 137.1-95.2 (qualitative ion), with collision energy set at 9 V and 13 V respectively. Cell accelerator voltage was set at 4 V, with dwell time at 100 ms.

GC-MS analysis of limonene. Limonene samples were diluted in hexane and run on an Agilent 7890B gas chromatography mass spectrometry (GC-MS) system with a DB-5 column (5%-phenyl-methylpolysiloxane, 30 m x 0.25 mm x 0.25 μm, Agilent Technologies, USA). For each run, 1 μL sample was injected with a 10:1 split ratio and 10 mL/min split flow. The GC oven temperature program was set at 40°C for 3 min, followed by a 10°C/min ramp to 100ºC and another 60°C/min ramp to 220°C with a hold time of 2 min. The injector and MS transfer line temperatures were at 250°C and 280°C, respectively. The MS was operated in selected ion-monitoring (SIM) mode using ions of m/z 136, 68 and 93, representing the molecular ion and two abundant fragmental ions of limonene. For chiral analysis, SPME vials were run through a CycloSil-B column (Agilent Technologies, USA). For each run, vials were agitated at 250 rpm, 60°C for 10 minutes. Samples were extracted for 10 minutes, followed by fiber bakeout at 270°C for 15 minutes. The GC oven temperature program was set at 80°C, followed by a 5°C/min ramp to 100°C, which was held for 3 minutes, follow by another 5°C/min ramp to 120°C, and a final 20°C/min ramp to 200°C, held for 3 minutes. The injector and MS transfer line temperatures were at 250°C. The MS was operated in selected scan mode, and ions of m/z 136, 68 and 93 were used for qualitative analysis.

HPLC Analysis of sugars and organic acids. Fermentation broth (1 mL) was centrifuged at 15,000 × g for 10 minutes to separate the supernatant from cells. The supernatant was collected and filtered through a 0.2 μm membrane filter to remove remaining particulates. Filtered

supernatant (200 µL) was used for the quantification of glycerol and acetate. Analysis was performed using a Waters Arc HPLC system equipped with a 2414 Refractive Index (RI) detector. The mobile phase consisted of 5 mM sulfuric acid, which was pumped through an Aminex HPX-87H column (Bio-Rad) at a flow rate of 0.6 mL/min. The column was maintained at 40°C during analysis. Retention times were compared to those of known standards for quantification.

Protein purifications

psY7 containing the gene of interest (tCsLS or tCsLS-Q8K) were transformed in *E. coli* BL21-Gold (DE3) competent cells by heat shock. Proteins were overexpressed in ZYM media (9.6 g/L tryptone, 4.8 g/L yeast extract, 20 mM sodium phosphate dibasic, 25 mM potassium dihydrogen phosphate, 50 mM ammonium chloride, 5 mM sodium sulphate, 0.05% glucose, 0.5% glycerol, 2 mM magnesium sulphate, 0.02 mM iron (III) chloride). A single colony was picked and incubated overnight in 1 mL LB supplemented with 100 µg/mL ampicillin. Pre-culture was re-inoculated in 50 mL ZYM media supplemented with 30 mM lactose and 100 µg/mL ampicillin at OD_{600} 0.1, and incubated at 30°C, 300 rpm for 24 hrs. Cells were harvested by centrifuging the culture at 3,900 rpm, 4°C for 45 mins. The cell pellet was suspended in 2.5 mL of His-lysis buffer (50 mM Tris/HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 1 mg/mL lysozyme (BioBasic), 20 U/mL DNaseI (NEB) and 5% glycerol). The mixture was frozen at -80°C overnight. It was thawed at room temperature for 30 minutes and re-frozen at -80°C for 1-2 hours, and the process was repeated twice, before thawing at room temperature to obtain the cell lysate. The lysate was centrifugated at 15,000 rpm at 4°C for 45 mins. The supernatant was added to 1 mL of Ni-NTA beads (Cube Biotech), which were equilibrated in His-binding buffer (50 mM Tris/HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 5% glycerol), and incubated at 4°C overnight with end-to-end mixing. The beads were washed with 40 mL His-binding buffer. An equal amount of protease

buffer (50 mM Tris/HCl pH 8.0, 20mM imidazole, 2 mM DTT, 5% glycerol), was added to the solution, as well as 30 µL (18 IU) of 3C protease, His, Human (Genscript). The solution was incubated at 4°C overnight with end-to-end mixing. The cleaved proteins were eluted with 3 x 0.5 mL His-binding buffer, assessed for purity with SDS-PAGE, and concentration was checked using a spectrophotometer, using calculated $E_{1\%}$. Freshly purified enzymes were kept at 4°C for a maximum of 4 days for kinetics experiments.

Kinetics experiments

The activities of purified limonene synthases were characterized using the EnzChek Pyrophosphate Assay Kit (ThermoFisher). GPP and NPP (Echelon Biosciences) were used as substrates. Each reaction was set up in a 200 µL system in a 96-well flat bottom plate (Greiner). Each reaction consisted of 1x reaction buffer, 0.2 mM 2-amino-6-mercapto-7-methylpurine ribonucleoside, 1 U/µL purine nucleoside phosphorylase, 0.03 U/µL inorganic pyrophosphatase, 20 mM $MgCl_2$, to which was added either pyrophosphate standards (0,10,20,30,40 µM) or a fixed concentration of limonene synthase (0.134 µM for tCsLS with GPP, and 0.154 µM for tCsLS with NPP, and tCsLS-Q8K for both GPP and NPP) and different concentrations of substrate (serial dilution from 80 µM to 5 µM, and one well without substrate). For each substrate, a no enzyme control was performed. All the components were incubated at 26°C for 10 minutes before adding the pyrophosphate or substrate. Then, the absorbance at 360 nm (A_{360}) was immediately measured, every 30 seconds for 1 hour. For standard curve, max A_{360} versus pyrophosphate concentration graph was plotted and slope of the linear fit was computed for the conversion of A_{360} to pyrophosphate produced in an enzymatic reaction. For enzymatic reactions, the background A_{360} values observed for controls without limonene synthase were subtracted to obtain net absorbance. A_{360} values were converted to pyrophosphate produced using the slope of the calibration curve.

The slope was calculated for the linear portion of the pyrophosphate vs time plot, yielding the reaction rate (V in μM/min) for each substrate concentration. Further, [Substrate] vs V was plotted, and the kinetic parameters were computed by fitting the data to the Michaelis-Menten equation using nonlinear regression in GraphPad Prism.. Note that the $MgCl_2$ can be supplemented as Mg^{2+} is required for limonene synthase activity, and 20 mM was reported to be optimal[28]. Each experiment was performed in duplicate.

Statistical analysis

Statistical analyses were performed using GraphPad Prism version 10.0.1 for Windows, GraphPad Software, Boston, Massachusetts USA, www.graphpad.com. All P-values showed result from unpaired t tests assuming Gaussian distribution, with ns, *, **, *** and **** correspond to $P > 0.05$, $P \leq 0.05$, $P \leq 0.01$, $P \leq 0.001$ and $P \leq 0.0001$, respectively.

AUTHOR INFORMATION

Author Contributions

C.P.M.S., S.B. and C.S. performed the experiments; O.J.S. and C.P.M.S performed the automation and screening with SPARROW; P.N. and Y.K. performed LCMS analysis; C.P.M.S. and X.C. analyzed the data; J.E. and I.A. performed bioinformatics analysis; X.C. and I.A designed the study; All authors contributed to the manuscript writing and review; All authors approved the final version of the manuscript.

Funding Sources

This work is funded by the EcoCTs project with the support of the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. We would like to acknowledge the funding support from Singapore Institute of Food and Biotechnology Innovation (SIFBI), and Central Research Fund (Applied / Translational Research) from A*STAR.

Conflict of interest: Singapore Application No.10202402813S, filed at the Intellectual Property Office of Singapore.

SUPPORTING INFORMATION

Figure S1. Comparison of (+)-Limonene Production Between Full-Length and Truncated CsLS.

Figure S2. Detailed limonene specific yield improvements of the best mutants identified from the SSM screening and of double mutants.

Figure S3. Chiral column chromatograms of different limonene synthases for GPP and NPP of the best mutants.

Figure S4. Kinetics analysis of tCsLS (WT) and tCsLS-Q8K (Q8K).

Figure S5. SDS-PAGE analysis of tCsLS and tCsLS-Q8K expression levels and densitometry analysis.

Figure S6. Two-plasmid system used for shake flask and bioreactor limonene production.

Figure S7. Improvement of specific limonene production using engineered MG1655 ispA S80F strain, different carbon source and the Q8K mutant.

Table S1. Plasmids used in this study.

Table S2. Bacterial strains used in this study.

Table S3. Primers Used for Site-Saturation Mutagenesis in Mutant Library Construction.

ACKNOWLEDGMENT

We thank Taek Soon Lee for the pJBEI-6409 plasmid, through Addgene plasmids #47048 and #50603. Authors would like to thank the regional computing meso center CALMIP (Grant P19013 and P22027) for granting access to High Performance Computing resources.

REFERENCES

1. Jongedijk, E. *et al.* Biotechnological production of limonene in microorganisms. *Appl Microbiol Biotechnol* 100, 2927–2938 (2016).

2. Kvittingen, L., Sjursnes, B. J. & Schmid, R. Limonene in Citrus: A String of Unchecked Literature Citings? *J Chem Educ* 98, 3600–3607 (2021).

3. Rinaldi, M. A., Ferraz, C. A. & Scrutton, N. S. Alternative metabolic pathways and strategies to high-titre terpenoid production in: Escherichia coli. *Natural Product Reports* vol. 39 90–118 Preprint at https://doi.org/10.1039/d1np00025j (2022).

4. Vieira, A. J., Beserra, F. P., Souza, M. C., Totti, B. M. & Rozza, A. L. Limonene: Aroma of innovation in health and disease. *Chem Biol Interact* 283, 97–106 (2018).

5. Eddin, L. B. *et al.* Neuroprotective potential of limonene and limonene containing natural products. *Molecules* vol. 26 Preprint at https://doi.org/10.3390/molecules26154535 (2021).

6. Ibáñez, M. D., Sanchez-Ballester, N. M. & Blázquez, M. A. Encapsulated limonene: A pleasant lemon-like aroma with promising application in the agri-food industry. A review. *Molecules* vol. 25 Preprint at https://doi.org/10.3390/molecules25112598 (2020).

7. Espina, L., Gelaw, T. K., de Lamo-Castellví, S., Pagán, R. & García-Gonzalo, D. Mechanism of Bacterial Inactivation by (+)-Limonene and Its Potential Use in Food Preservation Combined Processes. *PLoS One* 8, (2013).

8. Ravichandran, C., Badgujar, P. C., Gundev, P. & Upadhyay, A. Review of toxicological assessment of d-limonene, a food and cosmetics additive. *Food and Chemical Toxicology* 120, 668–680 (2018).

9. Vatankhah Lotfabadi, S., Mortazavi, S. A. & Yeganehzad, S. Study on the release and sensory perception of encapsulated d-limonene flavor in crystal rock candy using the time–intensity analysis and HS-GC/MS spectrometry. *Food Sci Nutr* 8, 933–941 (2020).

10. Kleybolte, M. M., Zainer, L., Liu, J. Y., Stockmann, P. N. & Winnacker, M. (+)-Limonene-Lactam: Synthesis of a Sustainable Monomer for Ring-Opening Polymerization to Novel, Biobased Polyamides. *Macromol Rapid Commun* 43, 2200185 (2022).

11. Durkin, A. *et al.* Scale-up and Sustainability Evaluation of Biopolymer Production from Citrus Waste Offering Carbon Capture and Utilisation Pathway. *ChemistryOpen* 8, 668–688 (2019).

12. Mendez-Perez, D. *et al.* Production of jet fuel precursor monoterpenoids from engineered Escherichia coli. *Biotechnol Bioeng* 114, 1703–1712 (2017).

13. Ciriminna, R., Lomeli-Rodriguez, M., Demma Carà, P., Lopez-Sanchez, J. A. & Pagliaro, M. Limonene: a versatile chemical of the bioeconomy. *Chemical Communications* 50, 15288–15296 (2014).

14. Johnson, T. J. *et al.* Producing next-generation biofuels from filamentous cyanobacteria: An economic feasibility analysis. *Algal Res* 20, 218–228 (2016).

15. Thomsett, M. R., Moore, J. C., Buchard, A., Stockman, R. A. & Howdle, S. M. New renewably-sourced polyesters from limonene-derived monomers. *Green Chemistry* 21, 149–156 (2019).

16. Paggiola, G. *et al.* Can bio-based chemicals meet demand? Global and regional case-study around citrus waste-derived limonene as a solvent for cleaning applications. *Biofuels, Bioproducts and Biorefining* 10, 686–698 (2016).

17. Liu, X., Gmitter, F. G., Grosser, J. W. & Wang, Y. Effects of rootstocks on the flavor quality of huanglongbing-affected sweet orange juices using targeted flavoromics strategy. *RSC Adv* 13, 5590–5599 (2023).

18. Nichkova, M. *et al.* Immunochemical Screening of Pesticides (Simazine and Cypermethrin) in Orange Oil. *J Agric Food Chem* 57, 5673–5679 (2009).

19. Moser, S. & Pichler, H. Identifying and engineering the ideal microbial terpenoid production host. *Applied Microbiology and Biotechnology* vol. 103 5501–5516 Preprint at https://doi.org/10.1007/s00253-019-09892-y (2019).

20. Alonso-Gutierrez, J. *et al.* Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production. *Metab Eng* 19, 33–41 (2013).

21. Alonso-Gutierrez, J. *et al.* Principal component analysis of proteomics (PCAP) as a tool to direct metabolic engineering. *Metab Eng* 28, 123–133 (2015).

22. Schilmiller, A. L. *et al. Monoterpenes in the Glandular Trichomes of Tomato Are Synthesized from a Neryl Diphosphate Precursor Rather than Geranyl Diphosphate*. www.pnas.org/cgi/content/full/.

23. Trapp, S. C. & Croteau, R. B. *Genomic Organization of Plant Terpene Synthases and Molecular Evolutionary Implications*. (2001).

24. Jo¨rg Bohlmann, J., Meyer-Gauen, G. & Croteau, R. *This Contribution Is Part of the Special Series of Inaugural Articles by Members of the Plant Terpenoid Synthases: Molecular Biology and Phylogenetic Analysis (Terpene Cyclaseisoprenoidsplant Defensegenetic Engineeringsecondary Metabolism)*. *Biochemistry National Academy of Sciences* vol. 95 www.pnas.org. (1998).

25. Del Terra, L. *et al.* Functional characterization of three Coffea arabica L. monoterpene synthases: Insights into the enzymatic machinery of coffee aroma. *Phytochemistry* 89, 6–14 (2013).

26. Byun-McKay, A. *et al.* Wound-induced terpene synthase gene expression in Sitka spruce that exhibit resistance or susceptibility to attack by the white pine weevil. *Plant Physiol* 140, 1009–1021 (2006).

27. Landmann, C. *et al.* Cloning and functional characterization of three terpene synthases from lavender (Lavandula angustifolia). *Arch Biochem Biophys* 465, 417–429 (2007).

28. Morehouse, B. R. *et al.* Functional and Structural Characterization of a (+)-Limonene Synthase from Citrus sinensis. *Biochemistry* 56, 1706–1715 (2017).

29. Kumar, R. P. *et al.* Structural Characterization of Early Michaelis Complexes in the Reaction Catalyzed by (+)-Limonene Synthase from Citrus sinensis Using Fluorinated Substrate Analogues. *Biochemistry* 56, 1716–1725 (2017).

30. Christianson, D. W. Structural and Chemical Biology of Terpenoid Cyclases. *Chemical Reviews* vol. 117 11570–11648 Preprint at https://doi.org/10.1021/acs.chemrev.7b00287 (2017).

31. Zhou, K. & Peters, R. J. Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants. *Phytochemistry* 70, 366–369 (2009).

32. Srividya, N., Lange, I. & Lange, B. M. Determinants of Enantiospecificity in Limonene Synthases. *Biochemistry* 59, 1661–1664 (2020).

33. Ignea, C. *et al.* Orthogonal monoterpenoid biosynthesis in yeast constructed on an isomeric substrate. *Nat Commun* 10, (2019).

34. Wang, X. *et al.* Enhanced limonene production in cyanobacteria reveals photosynthesis limitations. *Proc Natl Acad Sci U S A* 113, 14225–14230 (2016).

35. Rajaonarivony, J. I. M., Gershenzon, J. & Croteau3, R. *Characterization and Mechanism of (4S)-Limonene Synthase, A Monoterpene Cyclase from the Glandular Trichomes of Peppermint (/Ventha X Piperita)'. ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS* vol. 296 (1992).

36. Back, M., Chappell, K. & Noel, J. *Truncation of Limonene Synthase Preprotein Provides a Fully Active 'Pseudomature' Form of This Monoterpene Cyclase and Reveals the Function of the Amino-Terminal Arginine Pair †. Science* vol. 277 (1997).

37. Ignea, C. *et al.* Synthesis of 11-carbon terpenoids in yeast using protein and metabolic engineering. *Nat Chem Biol* 14, 1090–1098 (2018).

38. Leferink, N. G. H. *et al.* Experiment and Simulation Reveal How Mutations in Functional Plasticity Regions Guide Plant Monoterpene Synthase Product Outcome. *ACS Catal* 8, 3780–3791 (2018).

39. Schiff, W. H. & Oprian, D. D. Mutational Analysis of (+)-Limonene Synthase. *Biochemistry* 62, 2472–2479 (2023).

40. 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).

41. Rolf, J., Julsing, M. K., Rosenthal, K. & Lütz, S. A gram-scale limonene production process with engineered Escherichia coli. *Molecules* 25, (2020).

42. Sumbalova, L., Stourac, J., Martinek, T., Bednar, D. & Damborsky, J. HotSpot Wizard 3.0: Web server for automated design of mutations and smart libraries based on sequence input information. *Nucleic Acids Res* 46, W356–W362 (2018).

43. Lei, D., Qiu, Z., Qiao, J. & Zhao, G. R. Plasticity engineering of plant monoterpene synthases and application for microbial production of monoterpenoids. *Biotechnology for Biofuels* vol. 14 Preprint at https://doi.org/10.1186/s13068-021-01998-8 (2021).

44. Fuse-Murakami, T., Matsumoto, R. & Kanamori, T. N-Terminal Amino Acid Affects the Translation Efficiency at Lower Temperatures in a Reconstituted Protein Synthesis System. *Int J Mol Sci* 25, 5264 (2024).

45. Jia, B., Wang, T. & Lehmann, J. Peptidyl transferase center decompaction and structural constraints during early protein elongation on the ribosome. *Scientific Reports 2021 11:1* 11, 1–12 (2021).

46. Verma, M. *et al.* A short translational ramp determines the efficiency of protein synthesis. *Nature Communications 2019 10:1* 10, 1–15 (2019).

47. Zhao, W., Liu, S., Du, G. & Zhou, J. An efficient expression tag library based on self-assembling amphipathic peptides. *Microb Cell Fact* 18, 1–11 (2019).

48. Thompson, R. & Pickard, B. S. The amino acid composition of a protein influences its expression. *PLoS One* 19, e0284234 (2024).

49. Shukal, S., Lim, X. H., Zhang, C. & Chen, X. Metabolic engineering of Escherichia coli BL21 strain using simplified CRISPR-Cas9 and asymmetric homology arms recombineering. *Microb Cell Fact* 21, (2022).

50. Reiling, K. K. *et al.* Mono and diterpene production in Escherichia coli. *Biotechnol Bioeng* 87, 200–212 (2004).

51. Huang, C. N. *et al.* Mediating oxidative stress enhances α-ionone biosynthesis and strain robustness during process scaling up. *Microb Cell Fact* 21, (2022).

52. Shukal, S., Ong, L., Rehka, T., Chen, X. & Zhang, C. Microaerobic Fermentation Enables High-Titer Biosynthesis of the Rose Monoterpenes Geraniol and Geranyl Acetate in Escherichia coli. *ACS Sustain Chem Eng* 12, 3921–3932 (2024).

53. Chen, X. *et al.* Total enzymatic synthesis of cis-α-irone from a simple carbon source. *Nat Commun* 13, 7421 (2022).

For Table of Contents Use Only

Enhanced Production of (+)-Limonene through Targeted Engineering of *Citrus sinensis* Limonene Synthase

Clement P.M Scipion[1,2], Jérémy Esque[3], Shreyash Borkar[2], Ong Jing Sen[2], Cristalle Seah[1,2], Ong Leonard[2], Pnelope Ng[2], Yoganathan Kanagasundaram[2], Magali Remaud-Simeon[3], Sophie Bozonnet[3], Bo Xue[4,5,6] Wen Shan Yew[4.5.6], Isabelle André[3,], Xixian Chen[2,*]*

1. CNRS@CREATE, 1 Create Way, #08-01 Create Tower, Singapore 138602
2. Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), 31 Biopolis Way, Nanos, Singapore 138669
3. Toulouse Biotechnology Institute, TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France. 135, avenue de Rangueil, F-31077 Toulouse Cedex 04, France
4. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore (NUS), 8 Medical Drive, Singapore 117597, Singapore
5. NUS Synthetic Biology for Clinical and Technological Innovation, 28 Medical Drive, Singapore 117456, Singapore

6. Synthetic Biology Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, 14 Medical Drive, Singapore 117599, Singapore

Corresponding Authors

* Corresponding authors: isabelle.andre@insa-toulouse.fr; xixian_chen@sifbi.a-star.edu.sg