ACS Synthetic Biology

Research Article

pubs.acs.org/synthbio

Microbial Synthesis of Pinene

Stephen Sarria,[†] Betty Wong,[‡,§,#] Hector García Martín,[‡] Jay D. Keasling,[*,‡,§,∥] and Pamela Peralta-Yahya[*,†,‡,⊥]

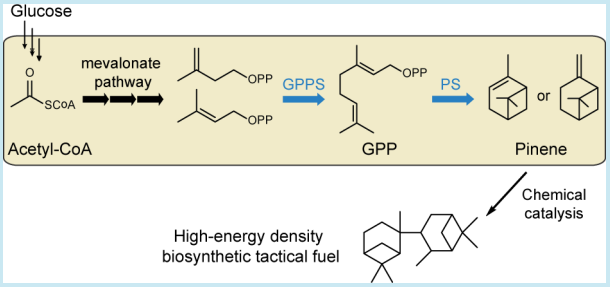
[†]School of Chemistry and Biochemistry and [⊥]School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

[‡]Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, 5885 Hollis Avenue, Emeryville, California 94608, United States

[§]QB3 Institute and [∥]Department of Chemical and Biomolecular Engineering and Department of Bioengineering, University of California, Berkeley, California 94720, United States

Supporting Information

ABSTRACT: The volumetric heating values of today's biofuels are too low to power energy-intensive aircraft, rockets, and missiles. Recently, pinene dimers were shown to have a volumetric heating value similar to that of the tactical fuel JP-10. To provide a sustainable source of pinene, we engineered *Escherichia coli* for pinene production. We combinatorially expressed three pinene synthases (PS) and three geranyl diphosphate synthases (GPPS), with the best combination achieving ∼28 mg/L of pinene. We speculated that pinene toxicity was limiting production; however, toxicity should not be limiting at current titers. Because GPPS is inhibited by geranyl diphosphate (GPP) and to increase flux through the pathway, we combinatorially constructed GPPS-PS protein fusions. The *Abies grandis* GPPS-PS fusion produced 32 mg/L of pinene, a 6-fold improvement over the highest titer previously reported in engineered *E. coli*. Finally, we investigated the pinene isomer ratio of our pinene-producing microbe and discovered that the isomer profile is determined not only by the identity of the PS used but also by the identity of the GPPS with which the PS is paired. We demonstrated that the GPP concentration available to PS for cyclization alters the pinene isomer ratio.



KEYWORDS: *advanced biofuels, tactical fuels, isoprenoids, terpene synthase, pinene*

Advanced biofuels have properties similar to petroleum-based fuels and can be "dropped in" to the existing transportation infrastructure.[1] Recent progress in engineering microbes for the production of advanced biofuels has resulted in biosynthetic alternatives to gasoline, such as butanol;[2] diesel, such as fatty acid ethyl esters;[3] and diesel precursors, such as bisabolene[4] and farnesene.[5] However, the development of microbial platforms for the production of high-energy density fuels, i.e., tactical fuels for use in aircraft and aircraft-launched missiles, has lagged behind. Current biosynthetic jet fuels, such as hydroprocessed esters and fatty acids (HEFA), derived from the natural oils present in oil-seed plants such as Camelina[6] and algae triglyceride,[7] have been used to power both military and commercial aircraft in 50:50 blends with Jet-A fuel.[8] More recently, dehydration of butanol into butene followed by oligomerization resulted in butene oligomers that can also be used as jet fuel.[9,10] Existing biosynthetic jet fuels, however, lack the volumetric energy content required to replace high-energy density fuels such as the tactical fuels JP-10, tetrahydrodicyclopentadiene, and RJ-5, a mixture of norbornadiene dimers used for aircraft-launched missiles (Figure 1). Attaining the volumetric energy content necessary for tactical fuels requires mimicking the strained ring systems found in JP-10 and RJ-5. Recently, pinene dimers have been show to contain high volumetric energy similar to that found in JP-10.[11] Pinene dimers are synthesized via chemical dimerization of pinene, a bicyclic terpene.[11]

Terpenes, such as pinene, are plant natural products with a wide range of functions from defense to pollinator attractants and allelopathy compounds.[12] More than 1,000 naturally occurring monoterpenes have been identified,[13] and monoterpenes play commercial roles as flavors, fragrances, insecticides, and pharmaceuticals.[14] Today, the major source of pinene is turpentine, a byproduct of the wood pulp industry.[15] In plants, monoterpenes (C10) such as pinene are biosynthesized in the plastid from the C5 intermediates isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) generated via the deoxyxylulose-5-phosphate (DXP) pathway. Geranyl diphosphate synthase (GPPS) carries out a head-to-tail condensation of IPP and DMAPP to produce geranyl diphosphate (GPP, C10), which is, in turn, cyclized by pinene synthase (PS) to produce either α- or β-pinene. In contrast, sesquiterpenes (C15), such as bisabolene, are produced via the mevalonate pathway in the cytosol. In the cytosol, IPP and DMAPP are condensed by farnesyl

Received: September 14, 2013
Published: February 27, 2014

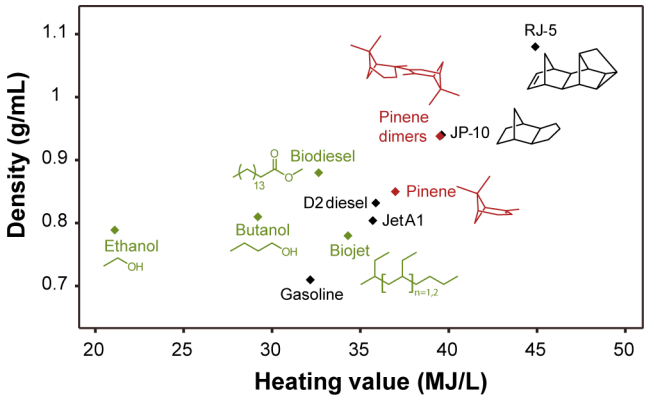


Figure 1. Energy density of petroleum-based fuels and advanced biofuels. Shown is the heating value of petroleum-based fuels (black) and advanced biofuels (green) as a function of density. Current advanced biofuels have lower density and heating value when compared to high-energy density petroleum-based fuels such as JP-10 and RJ-5. Pinene dimers (red) have density and heating value similar to that of JP-10. Pinene dimers mimic the strained ring systems found in JP-10 and RJ-5. Pinene dimers can be generated via pinene (red) dimerization using chemical catalysis. Data to generate this graph were obtained from the Biomass Energy Data Book 2011 and ref 11.

diphosphate synthase (FPPS) to produce farnesyl diphosphate (FPP, C15), which is, in turn, cyclized by sesquiterpene synthases into a variety of sesquiterpenes.

Given the large quantities of pinene dimers needed for use as a biofuel, engineering microorganisms to produce pinene from inexpensive sugars may be the most convenient and cost-effective approach to obtaining the necessary quantities of this advanced biofuel precursor. Using an engineered *Escherichia coli* strain for the overproduction of IPP and DMAPP via the mevalonate pathway,[16] we previously demonstrated pinene production of ∼1 mg/L from ionic liquid-treated switchgrass after introduction of a GPPS and PS.[17] More recently, *E. coli* was engineered to produce pinene at titers of ∼5 mg/L in shake flasks using a complex medium (beef broth) containing 2% glucose,[18] achieved in one of seven tested beef broths, the remainder of which resulted in ∼1 mg/L. Notably, microbial pinene titers are orders of magnitude lower than those of sesquiterpenes (bisabolene[4]) and diterpenes (taxadiene[19]). Further, a different monoterpene, limonene, has been produced microbially at a titer of ∼400 mg/L.[20] Likely reasons for the low microbial production of pinene may be (1) the toxicity of pinene or GPP to *E. coli*;[21] (2) the inhibition of GPPS by its substrate (GPP)[22] or high concentration of magnesium;[23,24] and (3) inhibition of PS by its substrate (GPP)[25] and product (pinene)[26] or reduced activity when magnesium is used as a cofactor for catalysis rather than manganese.[26]

Here, we engineered *E. coli* for the production of pinene, the immediate precursor to pinene dimers (Figure 2). Our strategy relied on combinatorially screening for high-flux PS and GPPS enzymes for the last two steps of the pathway. This was followed by protein fusion of GPPS to PS to reduce GPP

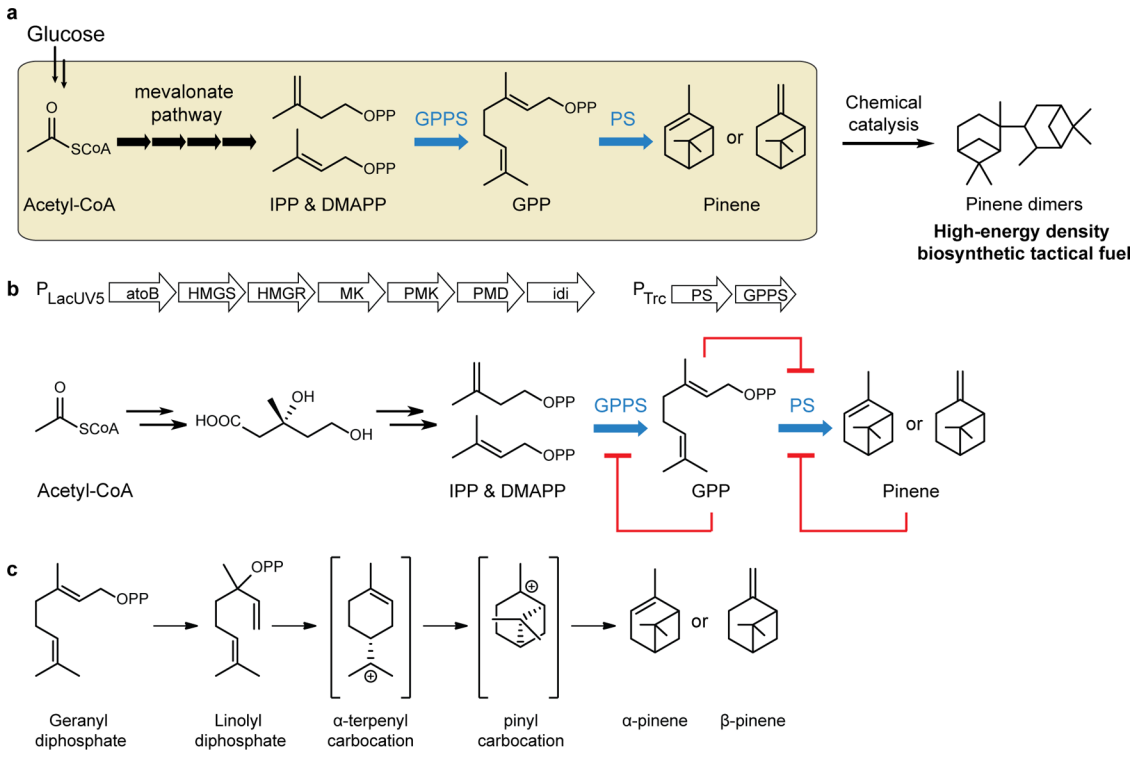


Figure 2. Microbial synthesis of pinene, the immediate precursor to a high-energy density biosynthetic tactical fuel. (a) *Escherichia coli* (yellow box) converts simple sugars into acetyl-CoA via primary metabolism. Introduction of a heterologous mevalonate pathway converts acetyl-CoA into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Addition of geranyl diphosphate synthase (GPPS) converts IPP and DMAPP into geranyl diphosphate (GPP), which is cyclized by pinene synthase (PS) to produce pinene. Pinene can be dimerized into pinene dimers via chemical catalysis to generate a high energy density biosynthetic tactical fuel. (b) Construct design. The mevalonate pathway is present in one plasmid as an operon under control of the LacUV5 promoter. The PS and GPPS are present in a separate plasmid as an operon under control of the Trc promoter. Also shown is the feedback regulation at the end of the pinene biosynthetic pathway. GPP inhibits GPPS and PS, while pinene inhibits PS. (c) Pinene synthase cyclization mechanism of GPP to pinene.

dx.doi.org/10.1021/sb4001382 | *ACS Synth. Biol.* 2014, 3, 466−475

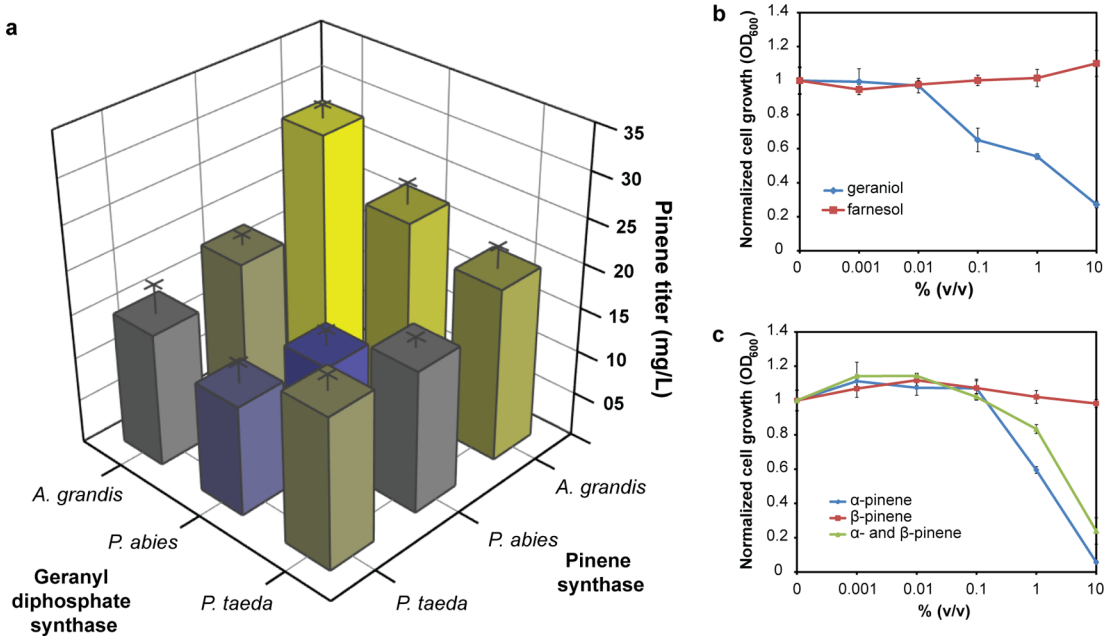


Figure 3. Microbial production of pinene via co-expression of geranyl diphosphate synthase and pinene synthase. (a) Pinene titers as a function of the GPPS and PS co-expressed. (b) Toxicity of geraniol, a proxy for GPP, and farnesol, a proxy for FPP, on cell growth. Cell growth was normalized to the growth of the same strain in the absence of geraniol or farnesol. (c) Toxicity of intracellularly produced GPP. Cell growth was normalized to the growth of the same strain not expressing GPPS. The experiments shown in panels a–c were done in triplicate, and the error bars represent the standard deviation from the mean.

inhibition of GPPS activity and to relieve potential GPP toxicity to the *E. coli* host. Using a previously engineered *E. coli* for the overproduction of IPP and DMAPP,[17] we combinatorially screened high-fidelity GPPS and PS enzymes from different plant conifers, resulting in strains that produced between 11.2 and 27.9 mg/L of pinene. To reduce GPP inhibition of GPPS activity, we constructed GPPS-PS protein fusions combinatorially, yielding pinene titers between 11.4 and 32.4 mg/L. Given the limited improvement obtained via GPPS-PS enzyme fusions, we speculated that the toxicity of pinene to the host was limiting production. We discovered that α-pinene is more toxic than β-pinene; however, pinene toxicity should not be limiting growth or pinene production at current titers. Finally, we investigated the pinene isomer ratio produced by our engineered *E. coli*, given the industrial desirability of β-pinene. We discovered that the pinene isomer profile is determined not only by the identity of the PS but also by the identity of the GPPS with which the PS is paired and whether the enzymes are co-expressed or fused. We also demonstrated that the GPP concentration available to PS for cyclization alters the pinene isomer ratio.

RESULTS AND DISCUSSION

Identifying Geranyl Diphosphate Synthases. Given that the majority of natural sources of pinene are plants, we examined plant GPPSs to construct the pinene-producing *E. coli* strain. Based on sequence analysis, three classes of plant GPPSs can be distinguished.[27] Heteromeric GPPSs, such as *Mentha piperita* GPPS,[28] are composed of a large and a small subunit and show no activity alone but together yield GPP and geranylgeranyl diphosphate (GGPP, C20) in the presence of excess IPP. There are two classes of homomeric GPPSs. One homomeric GPPS is found only in conifers, such as *Picea abies*[27] and *Abies grandis*,[23] and has high fidelity, producing only GPP. The other homomeric GPPS was first found in *Arabidopsis thaliana*,[29] and although it was initially identified as a GPPS, more recent data suggest that this enzyme may actually be a polyprenyl pyrophosphate synthase able to synthesize multiple products ranging from C25 to C45.[29] As we were interested in producing only GPP (C10), we chose only conifer homomeric GPPSs, in order to avoid the byproducts produced by the other GPPSs. Further, homomeric GPPS is potentially more likely to efficiently self-assemble than a heteromeric GPPS when expressed heterologously. Specifically, we screened *A. grandis*,[23] *P. abies*,[27] and *Pinus taeda* GPPS. We chose the GPPSs from *A. grandis* and *P. abies* because they are the best-characterized conifer GPPSs. We chose *P. taeda*, loblolly pine, as the third conifer GPPS source because it is a major source of turpentine in the United States. We hypothesized that loblolly pine could encode a highly efficient GPPS. As *P. taeda* GPPS has not been previously cloned, we assembled the gene from expressed sequence tags that have high similarity to conifer GPPS-like sequences. Finally, we decided against testing the previously engineered FPPS from *E. coli*, which produces GPP as 85% of its product profile (ISPA:S81F); although the engineered FPPS may have expressed well in our *E. coli* platform, it still produces FPP as a byproduct.[30]

Identifying Pinene Synthases. Our objective was to develop an *E. coli* platform for the production of pinene that, without any extra purification, could be used for the synthesis of pinene dimers. Therefore, we focused on high fidelity pinene synthases that cyclize GPP exclusively into α- or β-pinene. For use as biofuels, we have no preference with respect to which pinene isomer to produce, as both isomers can be transformed into pinene dimers.[11] For use as commodity chemicals, however, β-pinene is more valuable, as it is less abundant than α-pinene in turpentine.[15] From the 11 pinene synthases characterized to date,[31] we selected three pinene synthases with the highest fidelity (pinene synthases that produce pinene

dx.doi.org/10.1021/sb4001382 | ACS Synth. Biol. 2014, 3, 466–475

almost exclusively) for testing. Specifically, we selected *P. taeda* α-pinene synthase, which produces 100% α-pinene;[32] the mixed α/β pinene synthase from *A. grandis*, which produces ~42% α- and ~58% β-pinene;[33] and the mixed α/β pinene synthase from *P. abies*, which produces ~57% α- and ~27% β-pinene.[34] Conifer PSs, unlike sesquiterpene synthases or angiosperm PSs, prefer manganese over magnesium as a cofactor for catalysis.

Microbial Synthesis of Pinene via Coexpression. To identify the geranyl diphosphate/pinene synthase (GPPS/PS) pair that gives rise to the highest microbial production of pinene, we combinatorially screened the suite of three PSs and three GPPSs in an *E. coli* strain able to overproduce IPP and DMAPP. To overproduce IPP and DMAPP in *E. coli*, we introduced into *E. coli* MG1655 one vector carrying the genes from the mevalonate pathway to convert acetyl-CoA into IPP and DMAPP.[17] A second vector carried the genes for GPPS and PS in an operon driven by the strong IPTG-inducible promoter P_{Trc} (Figure 2b). Since GPPS and PS genes are endogenously expressed in plant plastids, we removed the plastid signal peptide and codon optimized their sequences to match the *E. coli* codon usage. Production of pinene by *E. coli* transformed with all combinations of the three PSs and three GPPSs was examined (Figure 3a). The GPPS/PS combination with the highest pinene titer was that from *A. grandis* GPPS/PS at 27.9 mg/L, while the lowest pinene titer at 11.2 mg/L resulted from the *P. abies* GPPS/PS pair. The same trend was found when comparing pinene specific production: *A. grandis* GPPS/PS led with 15.6 mg/L/OD_{600}, while *P. abies* GPPS/PS and *P. taeda* GPPS/PS had the lowest, at 3.0 mg/L/OD_{600} (Supplementary Figure SI1). The three microbial platforms with the highest pinene titers harbored the *A. grandis* PS, independent of the source of the GPPS used. Finally, although the cells expressing the *P. abies* GPPS/PS pair, the *P. taeda* GPPS/PS pair, and *A. grandis* GPPS/*P. taeda* PS pair grew to a higher final cell density than the other pinene-producing strains (Supplementary Figure SI2), all of the strains had similar specific growth rates (Supplementary Figure SI3).

To investigate whether PS or GPPS protein expression limited the performance of the pinene-producing microbes, we determined the protein levels of both soluble and insoluble PS and GPPS. With respect to PS, *P. abies* PS expressed the best in both the soluble and insoluble fraction. Remarkably, *A. grandis* PS resulted in the highest pinene titers even though the proteins in the soluble and insoluble fractions were barely visible on the Western blot (Supplementary Figure SI4). With respect to GPPS, *P. abies* and *A. grandis* GPPSs had similarly robust protein expression in the soluble and insoluble fractions and significantly higher protein expression than *P. taeda* GPPS (Supplementary Figure SI5). Interestingly, *P. taeda* GPPS resulted in pinene titers on par with *P. abies* GPPS, hinting at higher enzymatic activity or reduced inhibition of *P. taeda* GPPS by GPP.

Considering this information, we conclude that, in this system, *A. grandis* PS is the most efficient PS. Further, combinatorial screening of GPPS/PS pairs enabled us to reach pinene titers of ~28 mg/L using synthetic defined medium (EZ-rich) containing 1% glucose without medium optimization, which is a 5-fold increase over the previously reported 5.4 mg/L *E. coli* production of pinene in shake flasks[18] using complex medium containing 2% glucose and which required medium optimization. It is possible that pinene production is currently limited by (1) pinene or GPP toxicity to the microbial host, which may limit cell growth and further increases in pinene titers; (2) GPP inhibition of GPPS activity, resulting in low GPP pools for PS to act upon; or (3) diversion of GPP for the production of endogenous pyrophosphates, such as farnesyl pyrophosphate.

Geraniol Toxicity. As the same *E. coli* IPP and DMAPP overproduction strain has been previously used to produce the sesquiterpene bisabolene at ~400 mg/L,[4] the levels of IPP and DMAPP are not currently limiting pinene titers. We hypothesized that high intracellular GPP concentration may be toxic to the cell by, for example, inserting itself into the cell membrane. We measured GPP toxicity to *E. coli* in two ways: (1) by exogenous addition of geraniol, a proxy for GPP, to the microbial culture; and (2) by accumulating GPP intracellularly, via overexpression of GPPS. We used geraniol as a proxy for GPP because (1) both are C10 compounds and should inflict similar disruption on the cell membrane, due to their identical tails; (2) GPP would rapidly dephosphorylate when exogenously added to the medium resulting in geraniol; and (3) GPP may also dephosphorylate intracellularly to geraniol. Addition of geraniol to the medium results in reduced cell growth starting at 0.05% v/v, which translates to 445 mg/L of geraniol (Figure 3b). In contrast, farnesol, a proxy for the C15 diphosphate FPP and the intermediate to bisabolene,[4] shows no toxicity to *E. coli*. The concentration of geraniol required to observe reduced cell growth is currently 1 order of magnitude higher than the titers of microbially produced pinene. However, it is possible that internally produced GPP is more toxic than exogenously added geraniol. To test this hypothesis, we expressed only the GPPS in the IPP and DMAPP overproduction strain with the goal of accumulating GPP intracellularly. Interestingly, the specific growth rate of the IPP and DMAPP overproduction strain, in the absence of GPPS and expressing instead a terpene synthase to control for the cell burden of plasmid maintenance and protein expression (μ = 0.31 h^{-1}), was similar to that of the same strain expressing *A. grandis* GPPS (μ = 0.41 h^{-1}), *P. abies* GPPS (μ= 0.38 h^{-1}), or *P. taeda* GPPS (μ = 0.30 h^{-1}). Analysis of the cell culture revealed no geraniol produced in the strains expressing GPPS. We do note that, when measuring toxicity of intracellularly produced GPP, it is possible that, in the absence of PS, GPPS does not overproduce GPP, and therefore we do not see a detrimental effect of GPP. In summary, however, we cannot conclude that GPP toxicity to the cell host limits pinene production.

Pinene Toxicity. We measured the toxicity of both α- and β-pinene via exogenous addition and found that α-pinene is more toxic than β-pinene (Figure 3c). Specifically, α-pinene results in reduced cell growth starting at 0.5% v/v (~ 4.3 g/L) pinene in the medium. This concentration is higher than the one obtained using the best pinene-producing microbe (~28 mg/L). Cells harboring *P. taeda* PS, which preferentially produces α-pinene, should be the platform most highly impacted by α-pinene toxicity. However, cells harboring *P. taeda* PS had robust cell growth (Supplementary Figure SI2). Given that the two other PSs produce a mixture of α- and β-pinene, we determined the toxicity of a 50:50 mixture of α:β pinene. The α/β-pinene mixture reduced cell growth starting at 1% v/v (~ 8.6 g/L) pinene in the medium, significantly higher than that produced by our *A. grandis* GPPS-PS platform (~ 28 mg/L). Judging from these data, the titers produced by our engineered strains were not close to toxic pinene levels. However, toxicity could only be measured by exogenously

dx.doi.org/10.1021/sb4001382 | *ACS Synth. Biol.* 2014, 3, 466−475

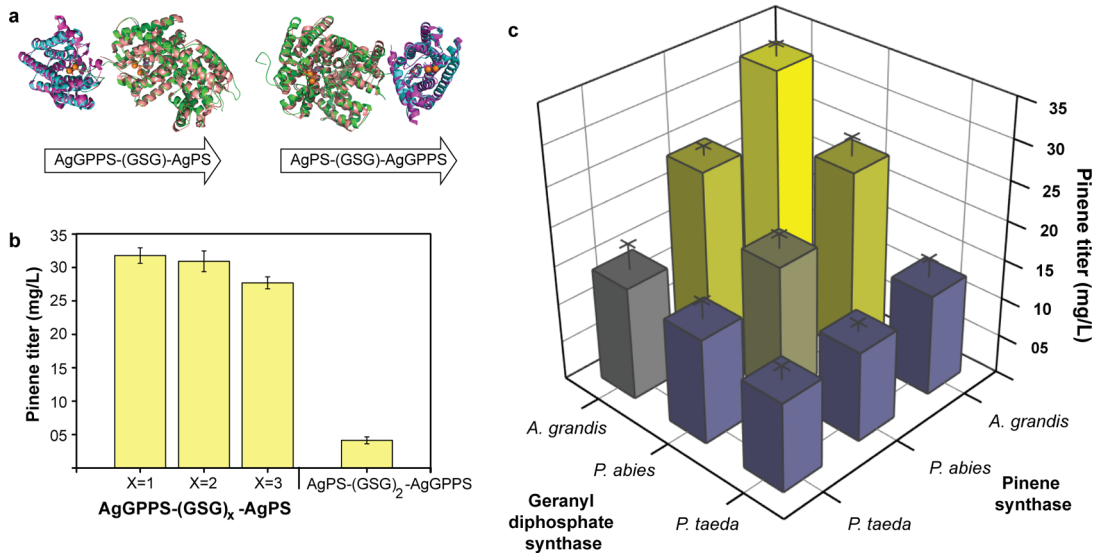


Figure 4. Microbial production of pinene via fusion of geranyl diphosphate synthase and pinene synthase. (a) Fusion protein design. GPPS at the N-terminus of the fusion results in GPPS and PS active sites facing one another. PS at the N-terminus of the fusion results in GPPS and PS facing in the same direction. Peach: *M. spicata* limonene synthase structure (PDB id: 2ONG). Magenta: *M. pipetira* GPPS (PDB id: 3OAC). Green and cyan: PS fused to GPPS. (b) Pinene titers of protein fusions with three, six, and nine amino acid linker lengths. (c) Microbial pinene titers as a function of the GPPS and PS protein fusions. The experiments shown in panels b and c were done in triplicate, and the error bars represent the standard deviation from the mean.

added pinene, and it is possible that internally produced pinene may be more toxic to the cell.

Overcoming GPPS Inhibition by GPP via Protein Fusions. To overcome potential GPPS inhibition by GPP, which may result in low GPP pools for PS to act upon, we envisioned generating protein fusions where the GPPS and PS active sites face one another so that GPP could be channeled from the GPPS active site directly into the PS active site. We used structural information to generate homology models of GPPS fused to PS in different orders, using a GSG linker, with the goal of determining the protein fusion directionality that would result in the active site of these enzymes facing one another. Specifically, using Phyre2,[35] we constructed structural homology models of *A. grandis* GPPS fused to *A. grandis* PS, the GPPS/PS enzyme pair leading to the best pinene titers. The homology model used *M. piperita* GPPS large subunit[28] to thread *A. grandis* GPPS and *Mentha spicata* limonene synthase[36] to thread *A. grandis* PS. Placing GPPS at the N-terminus of the fusion resulted in the PS and GPPS active sites facing one another, while placing PS at the N-terminus resulted in the PS and GPPS active sites facing in the same direction (Figure 4a). Therefore, we placed GPPS at the N-terminus and PS in the C-terminus of the protein fusion. To optimize the linker length between GPPS and PS, we tested linkers of three, six, and nine amino acids for the fusion (Figure 4b). The three- and six-amino acid linkers performed slightly better than the nine-amino acid linker. To ensure enough space for pinene release from the PS active site at the end of the reaction, we chose the six-amino acid linker length. To experimentally confirm that the directionality of the protein fusion was important, we constructed a protein fusion with PS at the N-terminus and GPPS at the C-terminus using the same six-amino acid linker. Constructing the protein fusion in this manner, with the active sites facing the same direction rather than each other, resulted in a pinene titer of 4 mg/L, which was even lower than the pinene titer obtained via co-expression of the same PS and GPPS. With this result in hand, we proceeded to test only fusions with the PS and GPPS active sites facing one another for increased activity.

Pinene Microbial Production Using Protein Fusions. Not knowing if *A. grandis* GPPS-$(GSG)_2$-PS would result in the highest pinene production, we constructed all possible combinations of GPPS-$(GSG)_2$-PS protein fusions, expressed them under the strong Trc promoter, and analyzed the resulting pinene production (Figure 4c). *A. grandis* GPPS-$(GSG)_2$-PS resulted in the highest pinene titer, at 32.4 mg/L, while *P. taeda* GPPS-$(GSG)_2$-PS resulted in the lowest pinene titer, at 11.4 mg/L. The *A. grandis* GPPS/PS fusion pair produced the most pinene in the co-expression (nonfusion) experiments as well (Figure 3a). Overall, five of the nine protein fusions showed improvement in pinene titers when compared to co-expression. The *P. abies* GPPS-$(GSG)_2$-PS fusion had the greatest improvement in pinene production, 52% when compared to co-expression, followed by the *A. grandis* GPPS-$(GSG)_2$-*P. abies* PS fusion, with a 35% improvement. The majority of constructs showed decreased pinene titers with the most significant drop in production coming from the *P. taeda* GPPS-$(GSG)_2$-*A. grandis* PS fusion, which dropped by 34%. The *A. grandis* GPPS-$(GSG)_2$-PS fusion led in the specific production of pinene, with 18.1 mg/L/OD_{600}, while the *P. taeda* GPPS-$(GSG)_2$-PS fusion had the lowest, at 1.8 mg/L/OD_{600}. (Supplementary Figure SI1).

All GPPSs and PSs investigated in this study have a conifer origin: *A. grandis*, *P. abies*, and *P. taeda*. We hypothesized that fusing enzymes from the same conifer species might result in higher pinene production than co-expressing them. Consistent with this hypothesis, we found that fusing the GPPS and PS from the same species in the case of *A. grandis* resulted in increased production over the co-expression. However, matching the GPPS and PS from the same species in the case of *P. taeda* worsened production.

Rationalization of Operon vs Protein Fusion Performance. Previous protein fusions of FPPS to C15 sesquiterpene synthases, bisabolene synthase and farnesene synthase, yielded

dx.doi.org/10.1021/sb4001382 | *ACS Synth. Biol.* 2014, 3, 466−475

2- and 8-fold improvements in sesquiterpene titers, respectively.[37,38] Protein fusions of geranylgeranyl diphosphate synthase to diacylglycerol diphosphate phosphatase resulted in ~2.5-fold improvement geranylgeraniol titers.[39] The marginal improvements seen in pinene titers when fusing GPPS to PS can be rationalized by the inhibition of PS by GPP. It is possible that, by bringing GPPS and PS together, the inhibition of GPPS by GPP is overcome but the inhibition of PS by GPP is exacerbated, resulting in a marginal overall improvement of pinene titers. To address this hypothesis, we determined the activity of PS in the presence of both manganese and magnesium in cell lysate at different GPP concentrations and confirmed that pinene synthesis decreases with increasing GPP concentration (Figure 5). Percent conversion of GPP to pinene

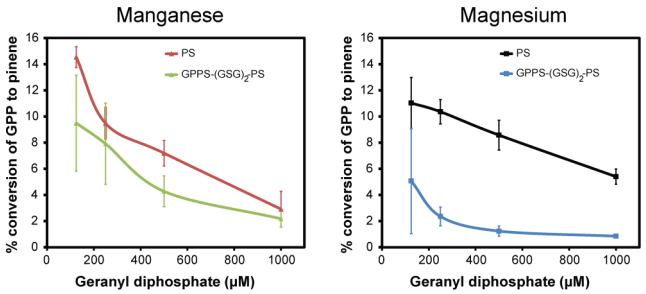


Figure 5. Analysis of GPP inhibition of pinene synthase alone and when part of a protein fusion with geranyl diphosphate synthase. (Left) Percent conversion of GPP to pinene in cell extract of *E. coli* expressing the *A. grandis* PS (PS) alone or the *A. grandis* geranyl diphosphate synthase-pinene synthase (AgGPPS-$(GSG)_2$-AgPS) protein fusion in the presence of manganese. (Right) Percent conversion of GPP to pinene in cell extract of *E. coli* expressing the *A. grandis* PS (PS) alone or the *A. grandis* geranyl diphosphate synthase-pinene synthase (AgGPPS-$(GSG)_2$-AgPS) protein fusion in the presence of magnesium. Percent conversion was calculated by dividing the moles of pinene produced over the moles of GPP at the beginning of the reaction. The experiments were done in triplicate, and the error bars represent the standard deviation from the mean.

by *A. grandis* PS decreased from 14.5% to 2.9% over the 0.1−1 mM range tested when manganese, the preferred cofactor, was used. The percent conversion decreased less sharply, from 11% to 5.4% over the same range, when using magnesium, which is the most likely cofactor used in *E. coli*. Next, we measured the activity of GPPS-$(GSG)_2$-PS protein fusion, in order to determine if the inhibition of PS was exacerbated when part of the fusion. In the presence of manganese, the *A. grandis* PS alone is 23% more inhibited than the *A. grandis* GPPS-PS protein fusion (Figure 5, left). In the presence of magnesium, however, the *A. grandis* GPPS-PS protein fusion is 33% more inhibited than the *A. grandis* PS alone (Figure 5, right). Given that magnesium is the most likely cofactor in *E. coli*, we conclude that inhibition of PS is exacerbated when it is part of a fusion in our system.

Ratios of α- and β-Pinene. Knowing that α-pinene was at least potentially toxic at high titers and noting that β-pinene is more expensive than α-pinene as it is less common in turpentine,[15] we investigated the α:β isomer ratios produced by the pinene-producing microbes. From the three pinene synthases, only *P. taeda* PS produced primarily α-pinene, while *A. grandis* PS and *P. abies* PS produced α/β-pinene mixtures. In the heterologous production of pinene, we expected that only the identity of the PS would determine the ratio of pinene isomers. However, we found that the identity of the GPPS and whether the GPPS/PS pair was a fusion affected the ratio of pinene isomers (Figure 6a). *P. taeda* PS consistently produced more α-pinene than β-pinene, independent of which GPPS it was paired with or whether it was co-expressed or part of a protein fusion. In contrast, the pinene isomer profile of *P. abies* PS and *A. grandis* PS changed depending on the identity of the GPPS with which they were paired. *A. grandis* PS resulted in a ~50:50 mixture of α- to β-pinene when paired with *A. grandis* GPPS. However, when paired with *P. taeda* GPPS, we observed a significant increase in α-pinene production, whether the enzymes were co-expressed or fused. The case with *P. abies* PS is more dramatic. Pairing *P. abies* PS with *P. taeda* GPPS resulted in mostly α-pinene production, while pairing it with *P. abies* GPPS as a co-expression resulted in mostly β-pinene and, as a fusion, resulted in a significant reduction of α-pinene.

In terpene synthases, the N-terminal strand caps the active site, shielding the carbocation intermediates from water.[36,40] Specifically, in monoterpene synthases, the N-terminus contains the $RR(X)_8W$ motif, where a pair of arginines stabilizes the holo conformation of the protein.[36] The pair of arginines has also been implicated in the isomerization of GPP to linalyl diphosphate, which is necessary to reach the α-terpenyl cation and, in turn, the pinyl carbocation and the final pinene structure[41] (Figure 2c). In the protein co-expression experiments, the PS cyclizes GPP into pinene without any movement limitation to the N-terminal strand; therefore, we do not expect any change in the PS cyclization mechanism. Rather, the only difference between the different co-expressions should be the level of GPP produced by the different GPPSs and thus available to the PS for catalysis. The GPP levels are determined by the origin of the GPPS (i.e., *A. grandis*, *P. taeda*, *P. abies*) with which the PS is co-expressed. Given the observed differences in α:β pinene ratios when *P. abies* PS and *A. grandis* PS are coupled to different GPPS, we hypothesize that GPP concentration regulates the product specificity of *P. abies* PS and *A. grandis* PS. To address this hypothesis, we measured the pinene isomer ratio at different GPP concentrations using the cell extract of *E. coli* expressing *A. grandis* PS, the PS with the greatest change in pinene isomer ratio *in vivo* (Figure 6b). Increasing the GPP concentration decreases the α:β pinene ratio from 1.76 at 1.25 μM GPP to 1.30 at 1 mM GPP when manganese is used as the *A. grandis* PS cofactor. Although manganese is the preferred *A. grandis* PS cofactor, the *E. coli* cytosol has at has at least 100-fold higher concentration of magnesium than manganese;[42] thus, we also tested the α:β pinene ratio in the presence of magnesium. Using magnesium, we also see a similar decrease in the α:β pinene ratio from 1.30 at 1.25 μM GPP to 0.98 at 1 mM GPP. PS isomer ratio regulation by GPPS may have been previously overlooked, as the PS product profile has been only measured at a single GPP concentration, and kinetic characterization of PS is traditionally carried out using radioactive or colorimetric assays.

In the protein fusion experiments, the movement of the PS N-terminal strand is restricted by its fusion to GPPS, which among other things may cause improper capping of the active site to water (Figure 6c). Additionally, the proximity of the N-terminal strand to the J-K loop, which stabilizes the holo form of the enzyme,[40] may explain the change in the pinene isomer ratio seen in the fusion proteins. Residues in the J-K loop have been previously shown to alter the terpene cyclization mechanism.[43] Specifically, a change in the pinene isomer ratio of *A. grandis* PS has been achieved via a single mutation at the

dx.doi.org/10.1021/sb4001382 | *ACS Synth. Biol.* 2014, 3, 466−475

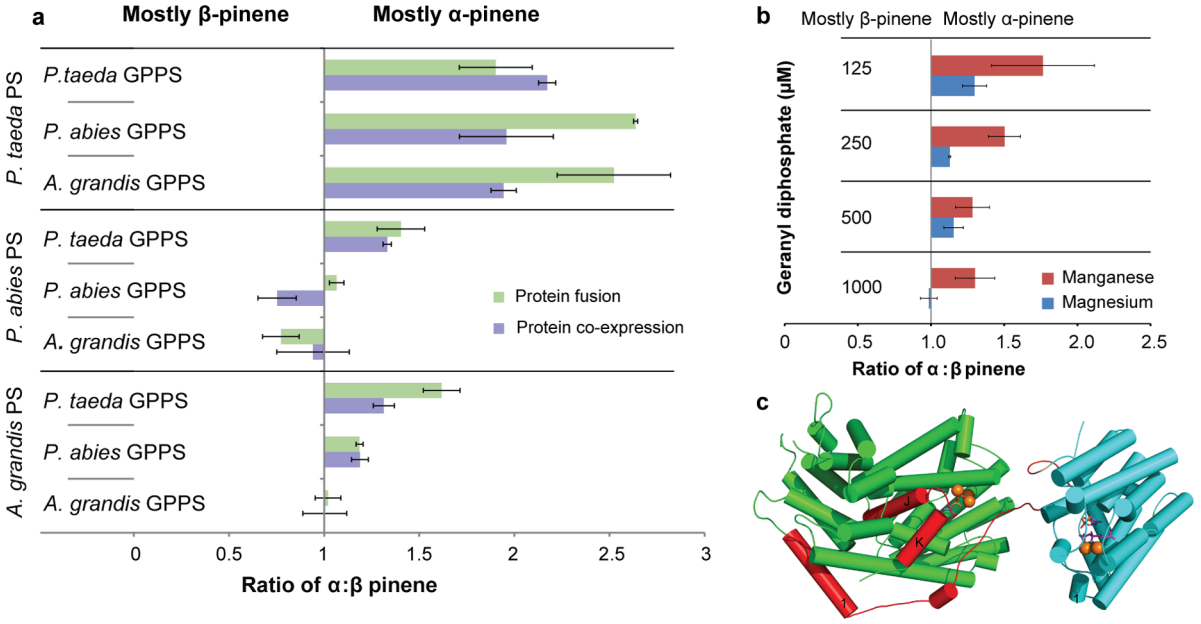


Figure 6. Microbially produced pinene isomer ratios as a function of geranyl diphosphate synthase (GPPS) and pinene synthase (PS). (a) Ratio of α:β-pinene when matching different GPPSs and PSs using either operons or fusions. (b) Ratio of α:β-pinene in cell extract of *E. coli* expressing *A. grandis* PS with either manganese or magnesium as the cofactor. (c) Homology model of *A. grandis* GPPS fused to *A. grandis* PS. The N-terminal strand and the J−K helix and loop of GPPS are highlighted in red. The experiments shown in panels a and b were done in triplicate, and the error bars represent the standard deviation from the mean.

end of the J-helix (F597W).[44] We hypothesize that the limited movement of the N-terminal strand hinders the proper localization of the J-K loop and, in turn, the J-helix, thus changing the pinene isomer ratios.

We have engineered *E. coli* for the renewable production of pinene, the immediate precursor to pinene dimers, a biosynthetic alternative to JP-10. By testing PSs and GPPSs from several organisms, we increased pinene production to 27.9 mg/L, 2-fold higher than the lowest pinene producing pair (11.2 mg/L). Combinatorial protein fusions of GPPS and PS resulted in slight improvement in pinene titers to ∼32 mg/L using *A. grandis* GPPS-$(GSG)_2$-PS. We rationalize the marginal improvement of pinene titers using protein fusions as due to PS inhibition by GPP, counterbalancing any improvements obtained by relief of GPPS inhibition by GPP. Therefore, we conclude that enzyme inhibition patterns must be taken into account to help determine the potential level of success of a protein fusion strategy. Given that the IPP- and DMAPP-overproducing *E. coli* strain used in this study has been previously used to achieve ∼400 mg/L of bisabolene, IPP and DMAPP precursor levels are not limiting; hence, the current bottleneck must lie in the last two enzymes in the pinene biosynthetic pathway. Therefore, we believe that further improvements in pinene titer will come from a combination of protein engineering approaches, such as generation of PSs with decreased GPPS inhibition, and better control of GPP intracellular levels.

Using flux balance analysis,[45] we calculated the theoretical yield of pinene produced from *E. coli* using the mevalonate pathway to be 0.270 g pinene/g glucose. Experimentally, we produced pinene at ∼1.2% of the pathway-dependent calculated theoretical yield, that is, assuming only glucose in the EZ-rich media is used for pinene production. Assuming a break-even price of glucose at the mill to be close to US $0.10/lb,[4] we calculate that the raw material cost of pinene production, ignoring non-sugar costs, to be approximately $68/kg of pinene. Assuming raw material costs to be only 50% of the final cost,[46] the final price of pinene would be $136/kg, or $443/gal, at current production levels. Assuming that we could produce pinene at the theoretical yield, the final price would be $1.63/kg, or $5.31/gal of pinene, resulting in a final price of ∼$6.42/gal for pinene dimers (assuming 90% yield[11] and negligible conversion cost), a significant savings over the current price of JP-10 of ∼$25/gal.[47] Commercial viability of microbial pinene for use as tactical fuel therefore requires reaching 26% theoretical yield.

Analysis of the pinene isomer profile of our microbial strain revealed that the ratios of α- to β-pinene vary not only due to the identity of the PS but also due to the identity of GPPS and whether the proteins are co-expressed or in a fusion. We demonstrated that the concentration of GPP available to *A. grandis* PS for catalysis alters the pinene isomer ratios, which explains the changes in isomer ratio seen in the protein co-expression experiments. A crystal structure of PS would aid in elucidating the mechanism of how GPP affects pinene synthase isomer ratio. Modulation of PS activity with GPP concentration may also occur *in planta* in response to stress. Indeed, changes in the α-pinene to β-pinene ratios have been previously seen after methyl jasmonate induction in pine.[48] We hypothesize that, in the GPPS-PS protein fusions, conformational restriction of the PS N-terminal strand may alter the holo conformation of PS, specifically the J-K loop, thus altering the terpene cyclization mechanism. *In vitro* biochemical analysis using purified PS-GPPS enzyme will be required to confirm the proposed changes in terpene cyclization mechanism.

Future engineering of microbial pinene production will require addressing the pathway problems at the enzyme level. Given that *A. grandis* PS resulted in the highest pinene production even though it was one of the more poorly expressed PSs, improving expression of this enzyme should

dx.doi.org/10.1021/sb4001382 | *ACS Synth. Biol.* 2014, 3, 466−475

result in increased pinene production. Alternatively, identifying GPPS and PS mutants not inhibited by GPP may also aid in increasing pinene titers. Finally, all of the PSs in this study prefer manganese as a cofactor over magnesium, which is far more available in the *E. coli* cytosol and, therefore, necessarily utilized in our experiments. Utilization of the magnesium may be limiting the activity of the PS *in vivo*; thus, identifying and expressing magnesium-dependent PSs may yield higher pinene titers as well.

METHODS

Construction of pPS-GPPS (Co-expressions). The nine PS and GPPS genes were codon optimized to match the *E. coli* codon usage and commercially synthesized (Genescript). For the pAgPS-XGPPS series, AgGPPS, PaGPPS, and PtGPPS were amplified from pAgGPPS, pPaGPPS, and pPtGPPS, respectively, using primers PPY61/PPY59 and cloned into pAgPS at *Bam*HI/*Hin*dIII. For the pPaPS-XGPPS series, AgGPPS, PaGPPS, and PtGPPS were amplified from pAgGPPS, pPaGPPS, and pPtGPPS, respectively, using primers PPY60/PPY59 and cloned into pPaPS at *Bam*HI/*Hin*dIII. For the pPtPS-XGPPS series, AgGPPS, PaGPPS, and PtGPPS were amplified from pAgGPPS, pPaGPPS, and pPtGPPS, respectively, using primers PPY58/PPY59 and cloned into pPtPS at *Bam*HI/*Hin*dIII.

Construction of pGPPS-$(GSG)_2$-PS (Protein Fusions). pAgGPPS-$(GSG)_2$-XPS series: To construct pAgGPPS-$(GSG)_2$-AgPS, AgGPPS was amplified from pAgGPPS using primers PPY90/BW33, and AgPS was amplified from ptAgPS using primers BW34/PPY91. Primer BW33 introduces the $(GSG)_2$ linker sequence. AgGPPS-$(GSG)_2$ and AgPS were fused using sewing PCR and cloned in pTRC99 at *Nco*I/*Hin*dIII. To construct pAgGPPS-$(GSG)_2$-PaPS, $(GSG)_2$-PaPS was amplified from pPaGPPS-$(GSG)_2$-PaPS using primers SS61/SS62 and cloned into pAgGPPS at *Xma*I/*Hin*dIII. To construct pAgGPPS-$(GSG)_2$-PtPS, PtPS was amplified from pPtPS using primers SS68/SS59 and cloned into pAgGPPS at *Xma*I/*Hin*dIII. Primer SS68 introduces the $(GSG)_2$ linker sequence. pPaGPPS-$(GSG)_2$-XPS series: To construct pPaGPPS-$(GSG)_2$-AgPS, $(GSG)_2$-AgPS was amplified from pAgGPPS-$(GSG)_2$-AgPS using primers SS45/SS46 and cloned into pPaGPPS at *Xma*I/*Hin*dIII. To construct pPaGPPS-$(GSG)_2$-PaPS, the PaPS leader sequence was amplified from pLead using primers SS35/SS36, and the truncated PaPS sequence was amplified from ptPaPS using primers SS37/SS38. Using SLIC[49] these two pieces were cloned into pPaGPPS at *Xma*I/*Hin*dIII. To construct pPaGPPS-$(GSG)_2$-PtPS, PtPS was amplified from pPtPS using primers SS69/SS59 and cloned into pPaGPPS at *Xma*I/*Hin*dIII. Primer SS69 introduces the $(GSG)_2$ linker sequence. pPtGPPS-$(GSG)_2$-XPS series: To construct pPtGPPS-$(GSG)_2$-AgPS, $(GSG)_2$-AgPS was amplified from pAgGPPS-$(GSG)_2$-AgPS using primers SS49/SS50 and cloned into pPtGPPS at *Xma*I/*Hin*dIII. To construct pPtGPPS-$(GSG)_2$-PaPS, $(GSG)_2$-PaPS was amplified from pPaGPPS-$(GSG)_2$-PaPS using primers SS63/SS62 and cloned into pPtGPPS at *Xma*I/*Hin*dIII. To constuct pPtGPPS-$(GSG)_2$-PtPS, the PtPS leader sequence was amplified from pLead using primers SS31/SS32, and the truncated PtPS sequence was amplified from ptPtPS using primers SS33/SS34. Using SLIC these two pieces were cloned into pPtGPPS at *Xma*I/*Hin*dIII.

Pinene Production and Quantification. *E. coli* MG1655 was co-transformed with pBbA5c-MevT-MBI containing the mevalonate pathway and plasmids containing GPPS and PS enzymes as operons or enzyme fusions. Cultures of strains co-transformed with both plasmids were grown in LB media overnight. Then, 100 μL of the overnight culture was used to inoculate 5 mL of EZ-rich medium (Teknova, 1% (v/v) glucose, amp^{100}, chl^{50}). The EZ-rich cultures were then incubated at 37 °C (250 rpm) until an OD_{600} of 0.8 was reached. Then, the cultures were induced with 1 mM IPTG and overlaid with 20% dodecane. After induction, the cultures were incubated at 30 °C (250 rpm) for 72 h. Then, 500 μL of the dodecane layer was placed in a 1.5-mL microcentrifuge tube and centrifuged at 25,000*g* for 1 min, and 50 μL of dodecane was diluted in 450 μL of ethyl acetate spiked with the internal standard limonene (10 μg/L). The samples were analyzed by GC/MS by using a standard curve of α and (−)-β-pinene (Sigma Aldrich). The GC/MS (Agilent 7890A with Agilent 5975 MS detector) was used with an Agilent DB-5MS column. The inlet temperature was set to 300 °C, flow at 1 mL/min, the oven at 50 °C for 30 s, ramp at 25 °C/min to 150 °C, and ramp of 40 °C/min to 250 °C.

Pinene and Geraniol Toxicity Measurements. An overnight culture of *E. coli* strain MG1655 was used to inoculate 5 mL of EZ-rich media (Teknova, 1% (v/v) glucose) containing varying concentrations of α, β, α+β-pinene, geraniol, or farnesol. After inoculation, the cultures were grown at 30 °C (250 rpm) for 24 h, and cell growth was measured as absorption at OD_{600}.

Pinene Conversion Using Cell Lysate. An overnight culture of *E. coli* MG1655 transformed with pAgPS or pAgGPPS-$(GSG)_2$-PS was used to inoculate 185 mL of LB medium at OD_{600} = 0.10 and incubated at 37 °C (250 rpm) until OD_{600} = 0.80 was reached. The culture was then induced with 1 mM IPTG and incubated at 30 °C (250 rpm) for 3 h. The culture was then split into two 90-mL cultures and centrifuged at 7354*g* for 20 min at 4 °C. Each pellet was resuspended in 9 mL of monoterpene synthase buffer[33] (50 mM Tris/HCl (pH = 7.5), 500 mM KCl, 5 mM dithiothreitol, 0.05% (w/v) $NaHSO_3$, and 10% (v/v) glycerol) with either 1 mM $MgCl_2$ or 1 mM $MnCl_2$ and sonicated using a Misonix Sonicator 3000 at 7.0 output level for 10 s, 30 s rest, for a total 4 min. After sonication, the lysate was centrifuged at 3220*g* for 30 min. Then, 980 μL of the supernatant was mixed with either 125 μM, 250 μM, 500 μM, or 1 mM GPP (Echelon) to a final volume of 1 mL, layered with 200 μL of dodecane, and incubated at 30 °C for 2 h. Pinene was quantified as described previously.

Flux Balance Analysis Yield Calculations. We used Flux Balance Analysis to determine the maximum yield for the mevalonate pathways for two different *E. coli* models (to provide extra robustness): iJR904[50] and iAF1260.[51] Using the COBRA package[52,53] we loaded each model and added the following reactions:

Pt30[c]: grdp ↔ appnn + ppi (pinene synthesis)

EXappnn(e): appnn[c] → (exchange reaction)

HMGCOAS[c]: coa + h + hmgcoa ↔ aacoa + accoa + h2o (mevalonate pathway)

HMGCOAR[c]: coa + mev-R + 2nadp ↔ 2h + hmgcoa + 2nadph

dx.doi.org/10.1021/sb4001382 | *ACS Synth. Biol.* 2014, 3, 466−475

MEVK1[c]: atp + mev-R → 5pmev + adp + h

PMEVK[c]: 5pmev + atp → 5dpmev + adp

DPMVD[c]: 5dpmev + atp → adp + co2 + ipdp + pi

We then knocked out DXPS (setting both lower bound and upper bound to zero), set glucose input (EX_glc(e)) to −6 mMol/gdw/h, changed the objective function to maximize EX_appnn(e), and solved for the maximum allowable value. The results for both models were multiplied by 136.23/180.16 (molecular weights of pinene and glucose, respectively) and divided by the incoming flux in order to obtain a gram of α-pinene per gram of glucose yield: 0.270 (iJR904) and 0.279 (iAF1266) g pinene/g glucose.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary methods, tables, figures, and sequences as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: jdkeasling@lbl.gov.

*E-mail: pperalta-yahya@chemistry.gatech.edu.

Present Address

[#]Biochemistry and Molecular Biophysics Program, California Institute of Technology, Pasadena, CA 91125.

Author Contributions

J.D.K. and P.P.-Y. conceived the project. S.S., B.W., and P.P.-Y. designed and performed the experiments. H.G.M. carried out the flux balance analysis. P.P.-Y, J.D.K., S.S, and H.G.M. wrote the manuscript.

Notes

The authors declare the following competing financial interest(s): J.D.K. has financial interest in Amyris, LS9, and Lygos.

■ ACKNOWLEDGMENTS

This work was started at the DOE Joint BioEnergy Institute (JBEI) and finished at the Georgia Institute of Technology. JBEI is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The work performed at the Georgia Institute of Technology was funded by Start-Up funds to P.P.-Y. and a Graduate Assistance in Areas of National Need (GAANN) fellowship to S.S. The authors thank Mario Ouellet for assembling the *Pinus taeda* geranyl diphosphate gene.

■ REFERENCES

(1) Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B., and Keasling, J. D. (2012) Microbial engineering for the production of advanced biofuels. *Nature 488*, 320−328.

(2) Bond-Watts, B. B., Bellerose, R. J., and Chang, M. C. Y. (2011) Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol. 7*, 222−227.

(3) Steen, E. J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., Del Cardayre, S. B., and Keasling, J. D. (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature 463*, 559−562.

(4) Peralta-Yahya, P. P., Ouellet, M., Chan, R., Mukhopadhyay, A., Keasling, J. D., and Lee, T. S. (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat. Commun. 2*, 483.

(5) Renniger, N., and McPhee, D. Fuel compositions comprising farnesane and farnesane derivatives and method of making and using same. Patent US20080098645 A1.

(6) Braukus, M. NASA Begins Flight Research Campaign Using Alternate Jet Fuel (http://www.nasa.gov/home/hqnews/2013/mar/HQ_13-066_ACCESS.html#.UjRmpD-bGk8) (Accessed, September 2013).

(7) Trimbur, D., Im, C.-S., Dillon, H., Day, A., Franklin, S., and Coragliotta, A. Renewable chemicals and fuels from oleginous yeast. Patent: US20110190522 A1.

(8) American Society for Testing and Materials (2013) D7566-13 Standard Specification for Aviation Turbine Fuel Containing Synthesized Hydrocarbons.

(9) Wright, M., Harvey, B., and Quintana, R. (2008) Highly efficient zirconium-catalyzed batch conversion of 1-butene: A new route to jet fuels. *Energy Fuels 22*, 3299−3302.

(10) Peters, M., and Taylor, J. Renewable jet fuel blendstock from isobutanol. Patent WO20111405610.

(11) Harvey, B., Benjamin, G., Wright, M., and Quintana, R. (2010) High-density renewable fuels based on the selective dimerization of pinenes. *Energy Fuels 24*, 267−273.

(12) Kirby, J., and Keasling, J. D. (2009) Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu. Rev. Plant Biol. 60*, 335−355.

(13) Seigler, D. (1998) *Plant Secondary Metabolism*, 2nd ed., Kluwer Academic Publisher, Norwell.

(14) Breitmaier, E. (2006) *Terpenes: Flavors, Frangances, Pharmaca, Pheromones*, Wiley-VCH, Tubingen.

(15) New Zealand Institute of Chemistry (2012) *Turpentine Production and Processing*. (http://nzic.org.nz/ChemProcesses/forestry/4F.pdf) (Accessed November 2013).

(16) Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat. Biotechnol. 21*, 796−802.

(17) Bokinsky, G., Peralta-Yahya, P. P., George, A., Holmes, B. M., Steen, E. J., Dietrich, J., Soon Lee, T., Tullman-Ercek, D., Voigt, C. A., Simmons, B. A., and Keasling, J. D. (2011) Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 108*, 19949−19954.

(18) Yang, J. M., Nie, Q. J., Ren, M., Feng, H. R., Jiang, X. L., Zheng, Y. N., Liu, M., Zhang, H. B., and Xian, M. (2013) Metabolic engineering of Escherichia coli for the biosynthesis of alpha-pinene. *Biotechnol. Biofuels 6*, 60.

(19) Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science 330*, 70−74.

(20) Alonso-Gutierrez, J., Chan, R., Batth, T. S., Adams, P. D., Keasling, J. D., Petzold, C. J., and Lee, T. S. (2013) Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production. *Metab. Eng. 19*, 33−41.

(21) Dunlop, M. J., Dossani, Z. Y., Szmidt, H. L., Chu, H. C., Lee, T. S., Keasling, J. D., Hadi, M. Z., and Mukhopadhyay, A. (2011) Engineering microbial biofuel tolerance and export using efflux pumps. *Mol. Syst. Biol. 7*, 487.

(22) Clastre, M., Bantignies, B., Feron, G., Soler, E., and Ambid, C. (1993) Purification and characterization of geranyl diphosphate synthase from Vitis vinifera L. cv Muscat de Frontignan cell cultures. *Plant Physiol. 102*, 205−211.

(23) Burke, C., and Croteau, R. (2002) Geranyl diphosphate synthase from Abies grandis: cDNA isolation, functional expression, and characterization. *Arch. Biochem. Biophys. 405*, 130−136.

(24) Tholl, D., Croteau, R., and Gershenzon, J. (2001) Partial purification and characterization of the short-chain prenyltransferases,

geranyl diphosphate synthase and farnesyl diphosphate synthase, from Abies grandis (grand fir). *Arch. Biochem. Biophys. 386*, 233−242.
(25) Phillips, M. A., Savage, T. J., and Croteau, R. (1999) Monoterpene synthases of loblolly pine (Pinus taeda) produce pinene isomers and enantiomers. *Arch. Biochem. Biophys. 372*, 197−204.
(26) Lewinsohn, E., Gijzen, M., and Croteau, R. (1992) Wound-inducible pinene cyclase from grand fir: purification, characterization, and renaturation after SDS-PAGE. *Arch. Biochem. Biophys. 293*, 167−173.
(27) Schmidt, A., and Gershenzon, J. (2008) Cloning and characterization of two different types of geranyl diphosphate synthases from Norway spruce (Picea abies). *Phytochemistry 69*, 49−57.
(28) Chang, T. H., Hsieh, F. L., Ko, T. P., Teng, K. H., Liang, P. H., and Wang, A. H. (2010) Structure of a heterotetrameric geranyl pyrophosphate synthase from mint (Mentha piperita) reveals intersubunit regulation. *Plant Cell 22*, 454−467.
(29) Hsieh, F. L., Chang, T. H., Ko, T. P., and Wang, A. H. (2011) Structure and mechanism of an Arabidopsis medium/long-chain-length prenyl pyrophosphate synthase. *Plant Physiol. 155*, 1079−1090.
(30) Reiling, K. K., Yoshikuni, Y., Martin, V. J., Newman, J., Bohlmann, J., and Keasling, J. D. (2004) Mono and diterpene production in Escherichia coli. *Biotechnol. Bioeng. 87*, 200−212.
(31) Degenhardt, J., Kollner, T. G., and Gershenzon, J. (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry 70*, 1621−1637.
(32) Phillips, M. A., Wildung, M. R., Williams, D. C., Hyatt, D. C., and Croteau, R. (2003) cDNA isolation, functional expression, and characterization of (+)-alpha-pinene synthase and (-)-alpha-pinene synthase from loblolly pine (Pinus taeda): stereocontrol in pinene biosynthesis. *Arch. Biochem. Biophys. 411*, 267−276.
(33) Bohlmann, J., Steele, C. L., and Croteau, R. (1997) Monoterpene synthases from grand fir (Abies grandis). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase. *J. Biol. Chem. 272*, 21784−21792.
(34) Martin, D. M., Faldt, J., and Bohlmann, J. (2004) Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol 135*, 1908−1927.
(35) Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc. 4*, 363−371.
(36) Hyatt, D. C., Youn, B., Zhao, Y., Santhamma, B., Coates, R. M., Croteau, R. B., and Kang, C. (2007) Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. *Proc. Natl. Acad. Sci. U.S.A. 104*, 5360−5365.
(37) Ozaydin, B., Burd, H., Lee, T. S., and Keasling, J. D. (2013) Carotenoid-based phenotypic screen of the yeast deletion collection reveals new genes with roles in isoprenoid production. *Metab. Eng. 15*, 174−183.
(38) Wang, C., Yoon, S. H., Jang, H. J., Chung, Y. R., Kim, J. Y., Choi, E. S., and Kim, S. W. (2011) Metabolic engineering of Escherichia coli for alpha-farnesene production. *Metab. Eng. 13*, 648−655.
(39) Tokuhiro, K., Muramatsu, M., Ohto, C., Kawaguchi, T., Obata, S., Muramoto, N., Hirai, M., Takahashi, H., Kondo, A., Sakuradani, E., and Shimizu, S. (2009) Overproduction of geranylgeraniol by metabolically engineered Saccharomyces cerevisiae. *Appl. Environ. Microbiol. 75*, 5536−5543.
(40) McAndrew, R. P., Peralta-Yahya, P. P., DeGiovanni, A., Pereira, J. H., Hadi, M. Z., Keasling, J. D., and Adams, P. D. (2011) Structure of a three-domain sesquiterpene synthase: a prospective target for advanced biofuels production. *Structure 19*, 1876−1884.
(41) Williams, D. C., McGarvey, D. J., Katahira, E. J., and Croteau, R. (1998) 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. *Biochemistry 37*, 12213−12220.
(42) Medicis, E. D., Paquette, J., Gauthier, J. J., and Shapcott, D. (1986) Magnesium and manganese content of halophilic bacteria. *Appl. Environ. Microbiol. 52*, 567−573.
(43) Yoshikuni, Y., Ferrin, T. E., and Keasling, J. D. (2006) Designed divergent evolution of enzyme function. *Nature 440*, 1078−1082.
(44) Hyatt, D. C., and Croteau, R. (2005) Mutational analysis of a monoterpene synthase reaction: altered catalysis through directed mutagenesis of (-)-pinene synthase from Abies grandis. *Arch. Biochem. Biophys. 439*, 222−233.
(45) Lewis, N. E., Nagarajan, H., and Palsson, B. O. (2012) Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat. Rev. Microbiol. 10*, 291−305.
(46) Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B. A., and Blanch, H. W. (2010) Technoeconomic analysis of biofuels: A wiki-based platform for lignocellulosic biorefineries. *Biomass Bioenerg. 34*, 1914−1921.
(47) Appeal of Raython Missile Systems, ASBCA No 5794, 11-2 B.C.A. 3854 (Armed Srv. Bd. of Cont. App. Oct. 11, 2011), available at www.asbca.mil/Decisions/2011/57594_101111_WEB.pdf
(48) Sampedro, L., Moreira, X., Llusia, J., Penuelas, J., and Zas, R. (2010) Genetics, phosphorus availability, and herbivore-derived induction as sources of phenotypic variation of leaf volatile terpenes in a pine species. *J. Exp. Bot. 61*, 4437−4447.
(49) Li, M. Z., and Elledge, S. J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods 4*, 251−256.
(50) Reed, J. L., Vo, T. D., Schilling, C. H., and Palsson, B. O. (2003) An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). *Genome Biol. 4*, R54.
(51) Feist, A. M., et al. (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol. 3*, 121.
(52) Becker, S. A., et al. (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc. 2*, 727−738.
(53) Schellenberger, J., et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc. 6*, 1290−1307.