Applied and Environmental Microbiology, Oct. 2007, p. 6277–6283 Vol. 73, No. 19
0099-2240/07/$08.00+0 doi:10.1128/AEM.00861-07

Identification of Isopentenol Biosynthetic Genes from *Bacillus subtilis* by a Screening Method Based on Isoprenoid Precursor Toxicity▽

Sydnor T. Withers,[1,2] Shayin S. Gottlieb,[3] Bonny Lieu,[3] Jack D. Newman,[3] and Jay D. Keasling[1,2,4,5]*

Department of Chemical Engineering, University of California, Berkeley, California[1]; Department of Bioengineering, University of California, Berkeley, California[2]; Amyris Biotechnologies, Emeryville, California[3]; Synthetic Biology Department, Physical Bioscience Division, Lawrence Berkeley National Laboratory, Berkeley, California[4]; and California Institute for Quantitative Biomedical Research (QB3), University of California, Berkeley, California[5]

Received 17 April 2007/Accepted 8 August 2007

We have developed a novel method to clone terpene synthase genes. This method relies on the inherent toxicity of the prenyl diphosphate precursors to terpenes, which resulted in a reduced-growth phenotype. When these precursors were consumed by a terpene synthase, normal growth was restored. We have demonstrated that this method is capable of enriching a population of engineered *Escherichia coli* for those clones that express the sesquiterpene-producing amorphadiene synthase. In addition, we enriched a library of genomic DNA from the isoprene-producing bacterium *Bacillus subtilis* strain 6051 in *E. coli* engineered to produce elevated levels of isopentenyl diphosphate and dimethylallyl diphosphate. The selection resulted in the discovery of two genes (*yhfR* and *nudF*) whose protein products acted directly on the prenyl diphosphate precursors and produced isopentenol. Expression of *nudF* in *E. coli* engineered with the mevalonate-based isopentenyl pyrophosphate biosynthetic pathway resulted in the production of isopentenol.

Isoprenoids have found use as neutraceuticals, flavors, fragrances, polymers, and drugs to treat malaria and cancer. Often, natural supplies are insufficient to support a commercial product due to ecological limitations. For example, the anti-inflammatory pseudopterosins harvested from gorgonian corals are used in the Estée Lauder product Resilience. The corals are harvested from Bahamian reefs, and demand exceeds the sustainable supply by approximately 5 to 1 (6). With little or no chance of supporting commercial production, promising lead compounds for serious diseases cannot be pursued in clinical trials under current FDA environmental impact requirements (7). Many of these limitations can be overcome though the microbial production of isoprenoids.

In order to produce these isoprenoids in a microbial host, the genes that code for the enzymes responsible for catalyzing the biosynthesis of these unique compounds must be cloned from their sources and expressed in the heterologous host. In particular, we need more genes for the terpene synthases that create the unique carbon skeletons of these compounds. The shortage of described terpene synthase genes has its origins in the available methods to discover new genes. Plant terpene synthase genes, particularly those found in the same organisms, have remarkably high levels of sequence similarity (2). Most methods to discover additional terpene synthases have taken advantage of this sequence similarity to isolate synthase genes from cDNA libraries (3). Thus, these methods are largely limited to plant terpene synthases, as there are very few data regarding sequence conservation outside of the plant kingdom. Additional genes have been discovered through whole-genome sequencing, as for *Arabidopsis thaliana* and *Mycobacterium tuberculosis* (2, 14). This is not yet a practical method for gene discovery, and it implicitly relies on sequence similarity to reveal the function of genes.

Here, we have devised a means to isolate terpene synthase clones from a population of bacterial clones expressing a DNA library. In prior work, we engineered a strain of *Escherichia coli* capable of producing elevated levels of terpene precursors from either acetyl coenzyme A (acetyl-CoA) or mevalonate (11). In this work, we took advantage of the toxicity of prenyl diphosphates and the ability to overcome that toxicity by functional expression of terpene synthases to screen a library of clones for those expressing a terpene synthase.

Plants are not the sole source of terpenes found in terrestrial environments. Some microorganisms, such as the soil bacterium *Streptomyces coelicolor*, produce geosmin, resulting in the commonly recognized "dirt" smell (8). The gram-positive bacterium *Bacillus subtilis* strain 6051 is known to produce perhaps the simplest of all isoprenoids, the hemiterpene isoprene (10). There is broad interest in isoprene synthesis, since isoprene is both the monomer unit for synthetic rubber and a potent greenhouse gas.

Isoprene synthases (IspS) have been isolated from kudzu and poplar, and they share many characteristics with other plant terpene synthases, implying a common evolutionary origin (12, 17) (Fig. 1). The poplar IspS is most similar (76% amino acid similarity) to the *Mentha spicata* limonene synthase, including the aspartate-rich motif that is highly conserved among terpene synthases. It is also known to utilize dimethylallyl diphosphate (DMAPP) as the substrate, with limited activity towards the monoterpene substrate geranyl diphosphate

* Corresponding author. Mailing address: Berkeley Center for Synthetic Biology, 717 Potter Street, Building 977, Mail Code 3224, University of California, Berkeley, CA 94720-3224. Phone: (510) 495-2620. Fax: (510) 495-2630. E-mail: keasling@berkeley.edu.
▽ Published ahead of print on 10 August 2007.

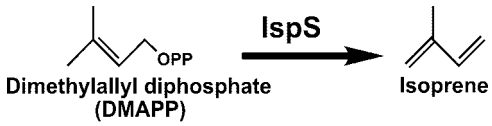

FIG. 1. Conversion of DMAPP to isoprene by IspS.

(12). No significant similarity exists between the known IspS genes and any gene in the *Bacillus subtilis* genome. The only isoprenoid genes that are apparent are those involved in central isoprenoid metabolism (9).

We sought to identify the gene responsible for isoprene biosynthesis in *B. subtilis* strain 6051 by use of a novel screening method based on overcoming the toxicity associated with the accumulation of prenyl diphosphate. We created a library of 19,000 clones harboring fragments of *Bacillus* genomic DNA. To screen for hemiterpene producers, we expressed the genes necessary to convert exogenous mevalonate to isopentenyl diphosphate (IPP) and DMAPP. In this strain background, we expressed the genes from the library. From the enriched library, we isolated two genes coding for proteins capable of overcoming the toxicity associated with accumulating IPP and DMAPP, *yhfR* and *nudF*. We have demonstrated that both protein products have an affinity for IPP and DMAPP, converting them into alcohol products. This work demonstrates the potential for this screening method to isolate genes involved in isoprenoid biosynthesis.

MATERIALS AND METHODS

Strains and plasmids. Four strains of *Escherichia coli* were used for these studies. *E. coli* DH10B [F^- *mcr*A Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lac*ZΔ*M15* Δ*lacX74 recA1 endA1 araD139* Δ(*ara leu*)*7697 galU galK* λ^- *rpsL nupG*] was used as the cloning and isoprenoid production strain. DH10B or DH10B-K1 was used for test enrichments.

E. coli DH10B-K1, a kanamycin-resistant strain derived from strain DH10B, was created to allow one to discriminate between it and DH10B in mixed culture studies. *E. coli* DH10B-K1 was created by transposon mutagenesis according to the manufacturer's directions (EZ-Tn5; Epicenter Biotechnologies). Electrocompetent DH10B cells were transformed with 1 μl of the supplied complex of the EZ-Tn5 <KAN2> transposon and transposase and plated on LB agar supplemented with 40 μg/ml kanamycin. Transformants were assayed for growth under conditions of toxicity and relief of toxicity by use of 96-well growth assays, as described below. DH10B-K1 showed the same phenotype as DH10B with regards to farnesyl diphosphate (FPP) toxicity, including relief of toxicity.

E. coli DH1 (F^- *supE44 hsdR17 recA1 gyrA96 relA1 endA1 thi-i1* λ^-) was used to express *nudF* and assay NudF activity.

E. coli W3110 [F^- *IN*(*rrnD-rrnE*)*1* λ^-] was used to screen for mutations in *lacI*. To do so, W3110 was grown on MacConkey lactose agar medium.

Table 1 summarizes the plasmids used in this study. Plasmids pMevT and pMBI contain the genes encoding the enzymes in the heterologous mevalonate-based IPP biosynthetic pathway (11). Expression of the genes harbored on these plasmids overproduces IPP and DMAPP. Plasmid pMBIS contains an additional gene, *ispA*, that allows the cell to synthesize FPP from IPP and DMAPP. Plasmid pTrc99A is a medium-copy-number expression plasmid with heterologous gene expression under the control of the hybrid *trc* promoter (1). Plasmid pADS contains the gene encoding amorpha-4,11-diene synthase cloned into pTrc99A (11).

Chemicals. DL-Mevalonolactone was purchased from Sigma-Aldrich (St. Louis, MO) and converted to DL-mevalonate by mixing 1 volume of 2 M DL-mevalonolactone with 1.02 volumes of 2 M KOH and incubating at 37°C for 30 min.

Competitive growth studies. *E. coli* DH10B and DH10B-K1 harboring the pMBIS plasmid were used for competitive growth comparisons. We found that the two strains were interchangeable with regard to apparent FPP toxicity. Strains harbored either pADS or pTrc99A. Overnight cultures of the two strains combined in various ratios from 1:1 to $1:10^5$ served as the inocula for 50 ml of medium supplemented with ampicillin (100 μg/ml) and tetracycline (10 μg/ml) in a 250-ml baffled flask. The medium was also supplemented with isopropyl-β-D-thiogalactopyranoside (IPTG) at 0.5 mM and mevalonate at 0 or 10 mM. At various times, a sample of the culture was withdrawn, diluted to 10^3 to 10^4 CFU/ml in LB medium, and plated on two different LB agar formulations, one supplemented with ampicillin and tetracycline and one supplemented with ampicillin and tetracycline plus kanamycin at 25 μg/ml. The sample optical density at 600 nm (OD_{600}) was also measured at this time on a Beckman DU-640 spectrophotometer. The resulting colonies were counted to determine the relative population levels in the culture.

Assay for growth inhibition. Growth was assayed in a 96-well spectrophotometer to test each individual strain's capacity to overcome growth inhibition due to prenyl diphosphate toxicity (SpectraMax; Molecular Devices). Various cultures were grown in 96-well plates with 100 μl of LB broth per well. Growth medium was supplemented with antibiotics, IPTG (0.5 mM), and various concentrations of mevalonate. Individual wells were inoculated with 2 μl of overnight culture diluted eightfold. Cultures were grown at 37°C with continuous shaking. The OD_{600} of each well was measured every 10 min. *E. coli* DH10B(pMBI) and DH10B(pMBIS) were tested for the relief of growth inhibition with pTrc99A as the negative control plasmid. pADS partially relieved toxicity in strains producing IPP and DMAPP from pMBI (data not shown), so this plasmid was used as a positive control for growth inhibition assays.

Library construction and screening. A library of genomic DNA from *Bacillus subtilis* strain 6051 was cloned into pTrc99A. Genomic DNA was purified from overnight cultures of *B. subtilis* 6051 by use of a QIAGEN tissue kit according to the manufacturer's instructions for gram-positive bacteria. The genomic DNA was partially digested with Sau3AI (New England Biolabs). DNA fragments of 1.5 to 4 kb were gel purified. The DNA was ligated into pTrc99A digested with BamHI and dephosphorylated with shrimp alkaline phosphatase (Promega). The library was transformed into electrocompetent *E. coli* DH10B. After 1 hour of outgrowth, 1 μl was plated to quantify the library size. The remaining outgrowth served as the inoculum for 100 ml of LB medium supplemented with 100 μg/ml ampicillin. This culture was shaken at 37°C overnight, and the plasmid DNA was

TABLE 1. Plasmids used in this study

Plasmid	Description	Antibiotic profile	Origin	Reference or source
pTrc99A	Expression vector under the control of the *trc* promoter	Amp^r	col E1	1
pADS	Synthetic ADS gene cloned into pTrc99A at NcoI-XmaI	Amp^r	col E1	11
pMevT	Single operon under *lac* control coding for acetoacetyl-CoA thiolase, HMG-CoA synthase, and HMG-CoA reductase	Cm^r	p15	11
pMBIS	pBBR1MCS-2 with the genes for mevalonate kinase, phosphomevalonate kinase, diphosphomevalonate decarboxylase, IPP isomerase, and FPP synthase cloned as a single operon	Tc^r	Bbr	11
pNudF	*nudF* cloned into pTrc99A at NcoI-XmaI		col E1	This work
pYhfR	*yhfR* cloned into pTrc99A at NcoI-XmaI		col E1	This work
pC9b	Genomic fragment from *B. subtilis* containing *nudF* cloned into pTrc99A at BamHI	Amp^r	col E1	This work

purified. The purified plasmid served as the starting DNA for subsequent library enrichment.

The genomic DNA library was transformed into the screening host, DH10B(pMBI). Libraries were also screened in W3110(pMBIS). Two hundred fifty microliters of the 1-ml transformation outgrowth was inoculated into 50 ml of LB medium supplemented with ampicillin (100 μg/ml) and tetracycline (10 μg/ml). This culture grew for 12 to 16 h at 37°C. One hundred twenty-five microliters of the overnight culture was inoculated into 50 ml of LB medium supplemented with antibiotics, IPTG (0.5 mM), and mevalonate. The library was enriched using two different concentrations of mevalonate, 5 and 10 mM. The cultures were shaken at 37°C for 9 h, and 125 μl of that culture was used to inoculate the next 50-ml enrichment culture. The enrichment continued for four rounds. The product of the fourth round was inoculated into a final 50-ml outgrowth culture of only LB medium with antibiotics and was shaken for 12 to 16 h at 37°C. Plasmid DNA was isolated using QIAGEN QIAprep miniprep kits, and the library DNA was gel purified away from the pMBI or pMBIS plasmid. The enriched library was subjected to an additional four rounds of enrichment or screened. Enriched libraries were transformed into *E. coli* W3110 and plated on MacConkey lactose solid agar medium. *E. coli* W3110 was chosen because this strain possesses a functional *lac* operon, unlike the other strains used in this study. Plasmids from red colonies, indicating a Lac^+ phenotype, were sequenced. Colonies harboring plasmids with mutant copies of *lacI*^{q} would remain white.

Cloning of *yhfR* and *nudF*. Both *yhfR* and *nudF* were subcloned into pTrc99A from library clones isolated after enrichment. The *yhfR* gene was amplified by PCR in a 100-μl reaction mixture containing 1× *Pfu* polymerase buffer [20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$ 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 0.1% mg/ml bovine serum albumin], 0.025 mM of each deoxynucleoside triphosphate [dNTP], 1 μM each of primers yhfR-F [ACAGCCGTTTGTTTAGTAAGACA] and yhfR-R [AGCACCCGGGATTATTTGATAAAGCCGGATAAG; the restriction site is underlined], 1 to 2 ng template plasmid, and 5 U *Pfu* polymerase [Stratagene]. The PCR program was 95°C for 1 min followed by 30 cycles of 95°C for 45 s, 60°C for 45 s, and 72° for 1 min followed by 72°C for 3 min. The ~600-bp PCR product was gel purified (QIAGEN), digested with XmaI, and phosphorylated with polynucleotide kinase. The *nudF* gene was amplified by PCR in a 100-μl reaction mixture containing 1× *Pfu* polymerase buffer, 0.025 mM of each dNTP, 1 μM each of primers nudF-F (AAATCATTAGAAGAAAAAACAATTGC) and nudF-R (ACGTCCCGGGATCATTTTTGTGCTTGGAGCG; the restriction site is underlined), 1 to 2 ng template plasmid, and 5 U *Pfu* polymerase (Stratagene). The PCR program was 95°C for 1 min followed by 30 cycles of 95°C for 45 s, 60°C for 45 s, and 72° for 1 min followed by 72°C for 3 min. The ~600-bp PCR product was gel purified (QIAGEN), digested with XmaI, and phosphorylated with polynucleotide kinase. Purified pTrc99A DNA was first digested with NcoI and purified with a PCR cleanup kit (QIAGEN). The overhanging ends were blunted with *Pfu* in a 55-μl reaction mixture incubated at 72°C for 5 min and containing 1× *Pfu* polymerase buffer, 48 μl purified DNA, 0.025 mM of each dNTP, and 2.5 U *Pfu* polymerase. The resulting DNA was again purified with a PCR cleanup kit and digested with XmaI and shrimp alkaline phosphatase. The previously described PCR products for *yhfR* and *nudF* were ligated into this vector to create pYhfR and pNudF. Plasmid pC9b was created by cloning a genomic fragment from an enriched library clone into a new pTrc99A backbone. Plasmid pC9 contained a region of *B. subtilis* 6051 genomic DNA including the *nudF* open reading frame (ORF). Plasmid pC9 was digested with KpnI and SalI, resulting in an ~1.1-kb DNA fragment that was ligated into pTrc99A, also digested with KpnI and SalI.

Heterologous expression of *nudF*. Plasmids pMevT and pMBI contain the genes encoding the enzymes in the heterologous mevalonate-based IPP biosynthetic pathway (11). Expression of the genes harbored on these plasmids overproduces IPP. Plasmids pMevT, pMBI, and pNudF were simultaneously transformed into chemically competent *E. coli* DH1. Three clones (transformants) were selected and grown overnight in 7 ml of LB medium with antibiotics at 37°C and then subcultured into 20 ml of LB medium with antibiotics to provide seed cultures for expression testing.

Growth and isopentenol production. *E. coli* DH1(pMevT)(pMBI)(pNudF) was grown in 250-ml baffled shake flasks containing 40 ml of 1× M9 salts (Fisher Scientific), 100 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) (J. T. Baker), 20 g/liter glucose, 5 g/liter yeast extract (Fisher Scientific), 100 μg/ml carbenicillin, 5 μg/ml tetracycline, and 34 μg/ml chloramphenicol. Seed cultures were subcultured into flasks to a final OD_{600} of 0.05. As a negative control, cultures containing only DH1 (no plasmids) were included. Cultures were incubated at 30°C and 250 rpm to an OD_{600} of 0.35 to 0.45 and then induced with IPTG (final concentration, 1 mM) and incubated at 30°C and 250 rpm for 66 hours. Samples were taken at regular intervals to quantify cell density and isopentenol concentration.

Lysate YhfR assays. DH10B harboring either pYhfR or the pTrc99A control were grown in 200 ml of LB medium supplemented with 50 μg/ml ampicillin to an OD_{600} of ~0.6, at which time IPTG was added to 0.5 mM. After 4 h, the cells were pelleted by centrifugation, and the pellets were stored at −80°C. The pellets were thawed and resuspended in 5 ml of buffer A [50 mM 3-(*N*-morpholinyl)-2-hydroxypropanesulfonic acid (MOPSO), pH 7.0, 20 mM $MgCl_2$, 2.5 mM $MnCl_2$, 1 mM dithiothreitol]. Water was added to a final volume of 9 ml. The cells were disrupted by sonication with a Virtis Virsonic sonicator for 3 cycles of 45 s at 67% duty and a power setting of 3.5. The lysates were clarified by centrifugation at 20,000 × *g* for 30 min. Protein concentrations were measured by Bradford assay (Bio-Rad). After clarification, 1.8 ml of lysate was dispensed into 10-ml serum bottles that were then sealed with septa. After equilibration to 37°C in a water bath, 200 μl of DMAPP diluted in water was added to the vials. The final DMAPP concentrations ranged from 0 to 200 mM. After incubating for 1 h, the liquid phase was sampled by exposing a solid-phase microextraction filament (polydimethylsiloxane, 85 μm by 1 cm; Supelco) for 2 min. The filaments were heated at 250°C for 2 min in the splitless inlet of an Agilent 6890/5973 gas chromatography (GC)-mass selective detector fitted with a Cyclosil-B column (30 m by 0.25 mm by 0.25 μm). The oven temperature profile was as follows: an initial 2-min period of 30°C followed by a 30°C/min ramp to 100°C and a final 120°C/min ramp to 300°C with a 2-min hold. The mass spectrometer was operated in selective ion-monitoring mode for ions 53, 67, and 68 *m/z*. Product identities were confirmed against authentic standards.

Quantification of isopentenol concentration. To extract isopentenol, 150 μl of culture broth was mixed with 750 μl of ethyl acetate for 15 min in a 1.8-ml GC vial. The mixture was allowed to settle, and 450 μl of the upper ethyl acetate phase was transferred to a new GC vial for analysis using an Agilent 6890N GC with an Agilent 6973MS and an HP-5MS column. The initial temperature was 60°C (3 min) followed by a 60°C/min ramp to 300°C for 2 min. The samples were compared to a commercial isopentenol standard.

RESULTS

Screening for sesquiterpene synthases. To demonstrate that the apparent toxicity of prenyl diphosphates could be exploited as a screen to identify terpene synthases, we tested the performance of a well-described sesquiterpene synthase (amorphadiene synthase [ADS]) in conditions of elevated intracellular FPP. The strains of *E. coli* used in these studies were engineered to produce elevated levels of prenyl diphosphate precursors to isoprenoids (11). Specifically, we utilized the pMBIS plasmid to convert mevalonate supplied in the medium to FPP. Prior work with this strain proved that expressing ADS converted the excess FPP to amorphadiene, while control strains harboring empty plasmid or expressing green fluorescent protein (GFP) accumulated FPP. This accumulation of FPP manifested itself as a reduced-growth phenotype in the control strains (11).

To test that the FPP toxicity would result in the takeover of a mixed population by the terpene synthase-expressing cells, we performed a series of experiments testing the relative growth rates of strains experiencing toxicity and of those not. In order to track the relative growth rates of two strains in mixed culture, we created a kanamycin-resistant derivative of *E. coli* DH10B, DH10B-K1. This strain performed identically to DH10B with respect to the toxicity phenotype when tested in 96-well plate growth assays (data not shown). Plating the mixed cultures on solid medium with and without kanamycin allowed us to determine the fraction of the population that was kanamycin resistant. Both strains were transformed with pMBIS to produce FPP from mevalonate. Additionally, DH10B-K1 was transformed with pADS to consume the FPP produced by pMBIS, and DH10B was transformed with pTrc99a as a control. The cultures were grown separately and then mixed at a 1:1 ratio, diluted into fresh LB medium con-

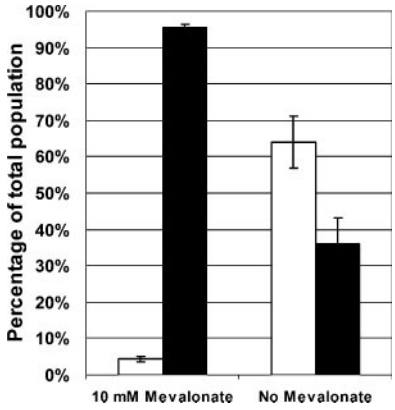

FIG. 2. Enrichment for ADS-expressing cells [DH10B(pMBIS)(pADS); black] over the empty vector control [DH10B-K1(pMBIS)(pTrc99A); white] in a mixed culture over 8.5 h. Cultures were inoculated at 0.5% with a 1:1 mixture of DH10B-K1(pMBIS)(pTrc99A) and DH10B(pMBIS)(pADS).

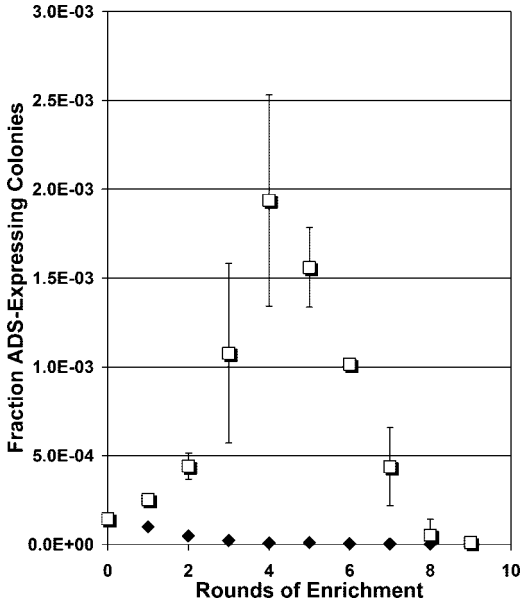

FIG. 3. Enrichment for and against ADS-expressing clones [DH10B-K1(pMBIS)(pADS)] in a 1:10,000 initial mixture with GFP-expressing clones [DH10B(pMBIS)(pTrcGFP)]. Cultures were tested in the presence (squares) and absence (diamonds) of 10 mM mevalonate.

taining 10 mM mevalonate, and grown. After 8.5 h of growth, samples were taken, and the proportion of kanamycin-resistant (ADS-containing) cells was quantified. The proportion of cells resistant to kanamycin was approximately 95%, indicating a strong selection for cells containing genes capable of transforming FPP (Fig. 2). Of interest are the actual colony counts for the pTrc99A-harboring cells. With the addition of mevalonate, the average number of viable colonies decreased from 600 ± 40 colonies to 21 ± 2. The ADS-expressing colony counts showed no significant difference from 350 ± 120 with no mevalonate to 470 ± 60 with 10 mM mevalonate. Errors given are 1 standard deviation from three replicates.

Additionally, pTrc99a was replaced with a vector expressing GFP (pTrcGFP) to control for protein expression effects and provide a second verification that the kanamycin plating was faithfully representing the populations. pTrcGFP was transformed into DH10B(pMBIS), and the transformed strain was mixed at a 1:10,000 ratio with DH10B-K1(pMBIS)(pADS). After four rounds of enrichment, DH10B-K1(pMBIS)(pADS) was enriched in the population by over an order of magnitude (Fig. 3). Beyond four rounds, there was a loss of pADS clones in the population to mutants. By use of W3110 in MacConkey agar, these mutations were shown to be plasmid-borne mutations in *lacI*^{q} (data not shown). Colonies harboring plasmids with the wild-type *lacI*^{q} were red, indicative of a Lac$^+$ phenotype, while mutant *lacI*^{q} expression resulted in a white colony phenotype, indicating a Lac$^-$ phenotype. Sequencing of mutant plasmids showed point mutations corresponding to the I^S type described by Suckow et al. (19). I^S mutants are characterized by unresponsiveness to inducer.

Library screening. This screening strategy was then used to screen a library of genomic DNA from the isoprene-producing bacterium *Bacillus subtilis* strain 6051 for the gene responsible for isoprene production. A library of genomic DNA with fragments in the 1.5- to 4-kb range was cloned into pTrc99A. Approximately 19,000 clones from this library were screened in *E. coli* DH10B harboring pMBI. The library was enriched in three replicate culture series. After four rounds of enrichment, the library was gel purified and transformed into *E. coli* strain W3110 to test for mutants in the plasmid-borne *lacI*. Ten clones from each replicate of the round four enrichment products were sequenced (Table 2).

Most of the clones were either *nudF* or *yhfR* clones. Fragments containing the same ORF were frequently found with different cloning junctions, indicating that the enrichment for *yhfR* and *nudF* was unlikely a result of cloning or transformation artifacts. Continued screening and sequencing resulted in additional clones with *yhfR* and *nudF* (data not shown). The annotation of the *Bacillus subtilis* genome lists *yhfR* and *nudF* as coding for 2,3-phosphoglycerate mutase and ADP-ribose pyrophosphatase, respectively. Additional screening using 5 mM mevalonate resulted in the same clones.

Effect of *yhfR* and *nudF* expression on toxicity phenotype. pYhfR was created to study the effects of YhfR on the toxicity phenotype. The *yhfR* gene was cloned into a fresh plasmid to eliminate the possibility of plasmid-borne mutations, such as those in *lacI* discussed earlier. The growth of *E. coli* DH10B(pMBI)(pYhfR) was assessed at various mevalonate concentrations (Fig. 4). Expression of YhfR relieved the toxicity associated with the overproduction of DMAPP and IPP. The growth of DH10B(pMBIS)(pYhfR) was also tested to determine the effect of YhfR expression on toxicity caused by the accumulation of FPP. YhfR expression had no effect on cells experiencing FPP-associated toxicity (data not shown), indicating that YhfR could transform IPP and/or DMAPP but not FPP.

Similar growth assays were performed on cultures expressing *nudF* from pNudF. DH10B(pMBI)(pNudF) grew poorly for all mevalonate concentrations studied. Since it was believed that the poor growth was related to the elevated NudF production due to the cloning of *nudF* directly behind the *trc* promoter with a strong ribosome binding site, the original fragment of

TABLE 2. Sequence data for 29 clones of the enriched *B. subtilis* 6051 library

Clone	Genome section[a]	Strand orientation[b]	Complete ORF(s)[c]	Start bp[d]
a1	2/21	+/+	*cypC*	16136
a2	18/21	+/+	*opuCB*, *A*	53538
a3	6/21	+/−	*yhfR*	98626
a4	6/21	+/−	*yhfR*	98654
a5	6/21	+/+	*yhfR*	96877
a6	13/21	+/+	*nudF*	48012
a7	13/21	+/+	*nudF*	48012
a8	10/21	+/+	*none*	155264
a9	6/21	+/+	*yhfR*	97154
a10	6/21	+/−	*yhfR*	98713
b1	13/21	+/+	*nudF*	48012
b2	6/21	+/+	*yhfR*	96877
b3	6/21	+/+	*yhfR*	96877
b4	6/21	+/+	*yhfR*	97129
b6	6/21	+/+	*yhfR*	96877
b7	6/21	+/+	*yhfR*	96877
b8	6/21	+/+	*yhfR*	96877
b9	6/21	+/−	*yhfR*	98713
b10	17/21	+/+	*yutG*, *F*, *E*, *D*	102420
c1	13/21	+/+	*nudF*	48012
c2	13/21	+/+	*nudF*	47444
c3	18/21	+/+	*chr*, *yvcK* and *L*	153828
c4	6/21	+/+	*yhfR*	96877
c5	9/21	+/−	*cheB*, *ylxH*	109331
c6	13/21	+/+	*nudF*	48012
c7	13/21	+/+	*nudF*	48012
c8	6/21	+/−	Fragments	136685
c9	13/21	+/−	*nudF*	49992
c10	14/21	+/+	*dnaJK*	11852

[a] Genome section numbers are for segments designated 1/21 to 21/21, corresponding to GenBank accession numbers Z99104 through Z99124, respectively; each segment represents 200 kb.

[b] The orientation of the cloned DNA with respect to the genome sequence is provided as +/+ or +/−. Sequences with +/+ orientation are cloned behind the *trc* promoter in the same direction as the reference genome sequence, whereas +/− oriented sequences are cloned in the reverse orientation.

[c] Annotated ORFs in the cloned region of the 5′ cloning junction. Note the dominance of *yhfR*- and *nudF*-containing clones.

[d] Base pair position of the 5′ cloning junction. The presence of multiple 5′ cloning sites indicates that *yhfR*- and *nudF*-containing clones arose independently from several genomic DNA library clones.

genomic DNA (containing the native promoter and terminator of *nudF* along with small fragments of other ORFs) was cloned into pTrc99A, creating plasmid pC9b. Using this construct, we were able to express NudF without any growth problems, most likely because this cloning orientation placed *nudF* under the control of the native *Bacillus* promoter rather than under the vector promoter. This is further supported by our sequence data from the enriched library (Table 2). We encountered genomic fragments containing *nudF* in both the plus and minus orientations relative to the plasmid *trc* promoter. When this reduced expression construct was tested in *E. coli* DH10B harboring pMBI, a normal growth phenotype was observed over a range of mevalonate concentrations (Fig. 5).

Functional expression in *E. coli*. To investigate further whether YhfR acted directly on DMAPP or IPP or if expression provided a general protective effect, we assayed for activity towards DMAPP. Lysates of DH10B(pYhfR) clearly show a product peak not present in the control; this peak was subsequently identified as 2-methyl-3-buten-1-ol (prenyl alcohol) by comparison with an authentic standard (Fig. 6).

While the pNudF constructs cloned directly behind the *trc* promoter showed very poor growth for all mevalonate concentrations studied with DH10B, only a small decrease in growth was seen when the same NudF construct was expressed in DH1 with the full mevalonate pathway. This pathway was contained on two plasmids, pMevT (containing *atoB* [acetoacetyl-CoA thiolase] and the 3-hydroxy-3-methyl-glutaryl [HMG]-CoA synthase and truncated HMG-CoA reductase genes) as described in the work of Martin et al. (11) and pMBI, and provided IPP as a substrate for pNudF (Fig. 7). Isopentenol production peaked at 43 h, reaching a final concentration of 1.3 mM. No isopentenol production was observed for the DH1 (−) control strain. Production lagged approximately 10 h behind growth in the experiment shown in Fig. 7. Even if protein production ceases with growth, continued product formation from existing enzymes would require only acetyl-CoA, NADPH, and ATP.

DISCUSSION

Much of the prior work identifying terpene synthases has relied heavily on the sequence similarity found among plant terpene synthases. Unfortunately, this method limits the scope

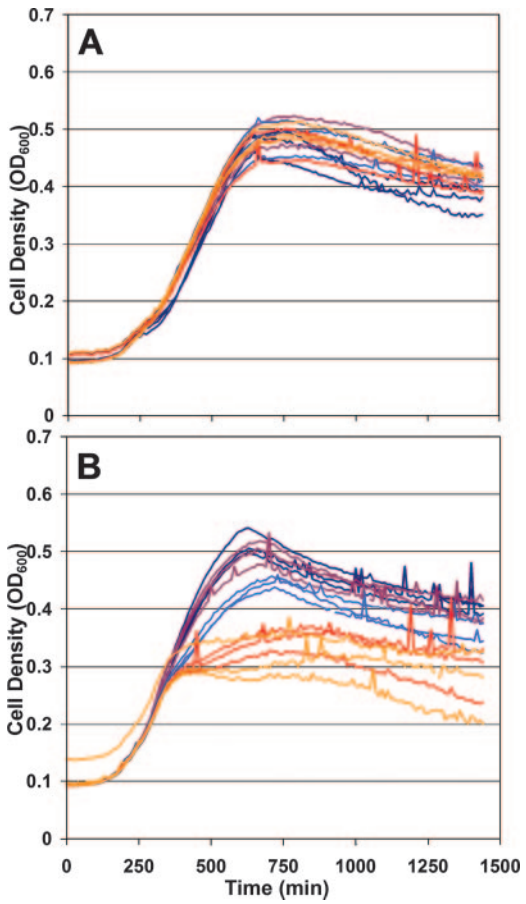

FIG. 4. YhfR expression is sufficient to overcome IPP/DMAPP toxicity. Growth of DH10B(pMBIS) with either pYhfR (A) or pTrc99A (B) in 96-well plates at various mevalonate concentrations: 0 mM (dark blue), 5 mM (purple), 10 mM (light blue), 15 mM (red), and 20 mM (yellow).

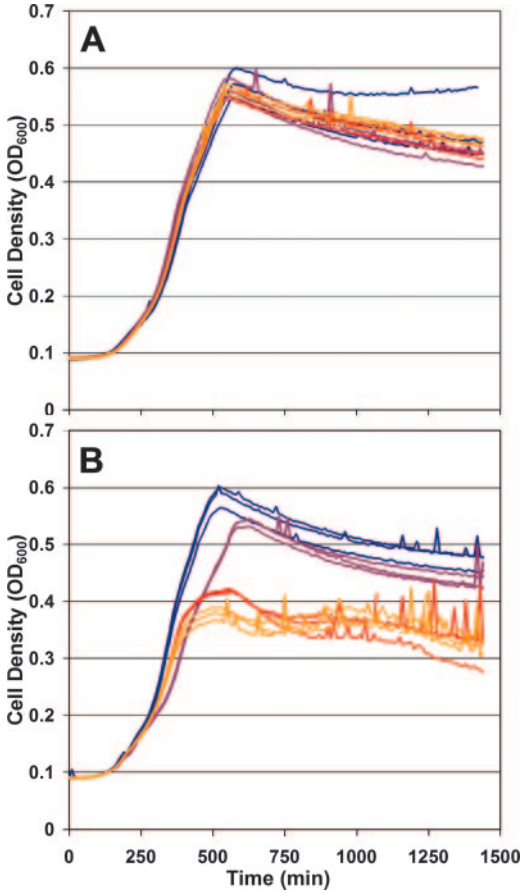

FIG. 5. Low-level NudF expression is sufficient to overcome IPP/DMAPP toxicity. Growth of DH10B(pMBIS) with either pC9b (A) or pTrc99A (B) in 96-well plates at various mevalonate concentrations: 0 mM (blue), 5 mM (purple), 15 mM (red), and 20 mM (yellow).

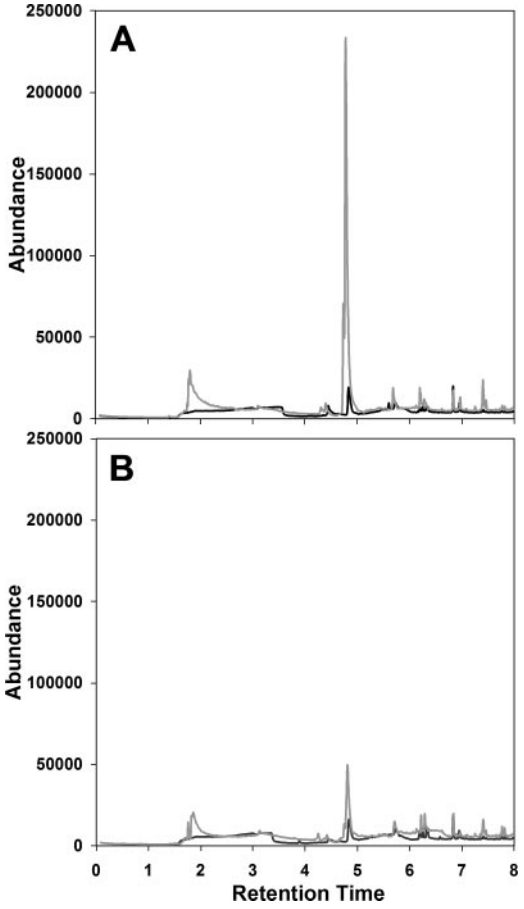

FIG. 6. Crude lysate assays with pYhfR (A) or pTrc99A (B) lysates with a DMAPP concentration of 0 (black) or 50 (gray) mM. The peak at a retention time of 4.9 min is 2-methyl-3-buten-1-ol.

of efforts to identify terpene synthases. We sought to overcome this limitation by developing a sequence-independent method to screen for the enzymatic consumption of prenyl diphosphates. This method is based on the toxicity of accumulated prenyl diphosphate precursors, first described by Martin et al. (11).

We chose to demonstrate the potential of this screen with plasmid mixtures including the ADS gene. This gene restored normal growth to cells accumulating FPP. In mixed culture experiments with ADS-expressing cells and control strains, we demonstrated that the ADS-expressing clones were able to overtake the populations. As shown in Fig. 2, a population of cells initially composed of equal parts of strains harboring either pADS or pTrc99A enriched to 96% pADS clones in the presence of mevalonate, representing a 4.5-fold enrichment. Of even more interest are the actual colony counts for the various strains. One notes that the pADS clones grew to nearly the same density with the addition of mevalonate. The enrichment was actually a product of the failure of the pTrc99A clones to grow. This method was expanded to serial enrichment of cultures to increase the total enrichment. Eventually, I^S-type mutations would arise in the plasmid-borne copy of *lacI*^{q}, resulting in a cell that expressed none of the genes necessary to produce the prenyl diphosphate toxicity. These cells expressing no heterologous proteins and experiencing no toxicity would have an advantage in a competitive growth situation.

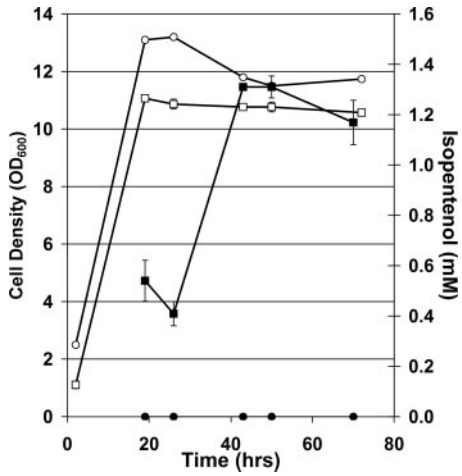

FIG. 7. Growth (open symbols) and isopentenol production (solid symbols) of *E. coli* DH1 expressing NudF (squares) or the pTrc99A control (circles).

Previous work has demonstrated that isoprene is produced by the DXP pathway in *Bacillus subtilis* strain 6051, yet no genes involved in isoprenoid secondary metabolism are apparent in the *Bacillus* genome (9, 18, 20). Sivy et al. isolated a fraction of *Bacillus* cell extract that catalyzed the conversion of DMAPP into isoprene (18). We sought to identify any proteins from *Bacillus* that when expressed could consume excess DMAPP and IPP in *E. coli*, and we screened an expression library of genomic DNA for this purpose. Clones carrying two genes, *yhfR* and *nudF*, dominated the population after four rounds of growth under conditions that produced lethal levels of DMAPP and IPP. Subsequent examination of growth under these same conditions showed that the corresponding proteins did restore a normal growth phenotype.

We demonstrated that both YhfR and NudF showed activity against prenyl diphosphates. Lysate assays indicated that YhfR converted DMAPP into prenyl alcohol. While YhfR is annotated as a cofactor-dependent 2,3-phosphoglycerate mutase, the purified protein does not exhibit mutase activity (15). The *Bacillus* sequence has both cofactor-dependent and cofactor-independent mutases in the genome. The deletion of *yhfR* results in no negative growth phenotype, but the deletion of the cofactor-independent mutase is lethal (15). The ortholog to YhfR in *Bacillus stearothermophilus*, PhoE, has been shown to be a general phosphatase (16). The genomic context of *yhfR* provides no insight into the purpose of the gene to *B. subtilis*. It does not appear to be transcribed with any other genes.

Expression in *E. coli* demonstrated that NudF has activity towards prenyl diphosphates similar to that of YhfR. As shown in Fig. 7, the production of isopentenol lagged behind growth by approximately 10 h. Even as protein production ceased with growth, continued product formation from existing enzymes required only acetyl-CoA, NADPH, and ATP. This is a likely scenario with the carbon-rich medium used in this study. NudF belongs to the Nudix hydrolase superfamily of enzymes, which typically catalyzes reactions on nucleoside diphosphates that are part of larger molecules. Specifically, NudF belongs to the ADP-ribose pyrophosphatase subfamily, and this activity has been verified experimentally (5). Its role in consuming DMAPP is unclear at this time.

In summary, the expansion of isoprenoid research requires that new methods be developed to identify genes from sources other than plants. Our work provides a sequence-independent method for identifying terpene synthase genes from libraries of DNA. This is underscored by the nature of the two *Bacillus* genes isolated, *yhfR* and *nudF*. Neither of the sequences offers any indication of involvement with terpene metabolism. This method has particular application where little information is available about the host organism, especially for complex assemblages of microbes. Just such assemblages have been identified as harboring the producing organism of isoprenoids in the gorgonian corals *Pseudopterogorgia elisabethae* and *P. bipinnata* (4, 13). This screen could also be adjusted for use in directed evolution studies performed on terpene synthases. Successful directed evolution studies require the ability to screen very large libraries ($>10^6$ clones) of mostly nonfunctional clones. One would need to adjust the mevalonate concentration and induction levels to produce prenyl diphosphates at a rate that just exceeds the capacity of the parent enzyme. This method would allow one to devote analytical resources only to clones with demonstrated activity.

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (BES-9911463), the Office of Naval Research (FDN00014-99-0182), and the University of California Discovery Grant Program (99-10044).

REFERENCES

1. **Amann, E., B. Ochs, and K. J. Abel.** 1988. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. Gene **69:**301–315.
2. **Aubourg, S., A. Lecharny, and J. Bohlmann.** 2002. Genomic analysis of the terpenoid synthase (AtTPS) gene family of Arabidopsis thaliana. Mol. Genet. Genomics **267:**730–745.
3. **Back, K., and J. Chappell.** 1995. Cloning and bacterial expression of a sesquiterpene cyclase from Hyoscyamus muticus and its molecular comparison to related terpene cyclases. J. Biol. Chem. **270:**7375–7381.
4. **Boehnlein, J. M., L. Z. Santiago-Vazquez, and R. G. Kerr.** 2005. Diterpene biosynthesis by the dinoflagellate symbiont of the Caribbean gorgonian Pseudopterogorgia bipinnata. Mar. Ecol. Prog. Ser. **303:**105–111.
5. **Dunn, C. A., S. F. O'Handley, D. N. Frick, and M. J. Bessman.** 1999. Studies on the ADP-ribose pyrophosphatase subfamily of the nudix hydrolases and tentative identification of trgB, a gene associated with tellurite resistance. J. Biol. Chem. **274:**32318–32324.
6. **Faulkner, D. J.** 2000. Marine pharmacology. Antonie Leeuwenhoek **77:**135–145.
7. **FDA.** 1998. Guidance for industry. Environmental assessment of human drugs and biologics applications. CDER/CBER CMC 6. FDA, Washington, DC.
8. **Gust, B., G. L. Challis, K. Fowler, T. Kieser, and K. F. Chater.** 2003. PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc. Natl. Acad. Sci. USA **100:**1541–1546.
9. **Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin, et al.** 1997. The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature **390:**249–256.
10. **Kuzma, J., M. Nemecek-Marshall, W. H. Pollock, and R. Fall.** 1995. Bacteria produce the volatile hydrocarbon isoprene. Curr. Microbiol. **30:**97–103.
11. **Martin, V. J., D. J. Pitera, S. T. Withers, J. D. Newman, and J. D. Keasling.** 2003. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat. Biotechnol. **21:**796–802.
12. **Miller, B., C. Oschinski, and W. Zimmer.** 2001. First isolation of an isoprene synthase gene from poplar and successful expression of the gene in Escherichia coli. Planta **213:**483–487.
13. **Mydlarz, L. D., R. S. Jacobs, J. Boehnlein, and R. G. Kerr.** 2003. Pseudopterosin biosynthesis in Symbiodinium sp., the dinoflagellate symbiont of Pseudopterogorgia elisabethae. Chem. Biol. **10:**1051–1056.
14. **Nakano, C., T. Okamura, T. Sato, T. Dairi, and T. Hoshino.** 2005. Mycobacterium tuberculosis H37Rv3377c encodes the diterpene cyclase for producing the halimane skeleton. Chem. Commun. (Cambridge) **2005:**1016–1018.
15. **Pearson, C. L., C. A. Loshon, L. B. Pedersen, B. Setlow, and P. Setlow.** 2000. Analysis of the function of a putative 2,3-diphosphoglyceric acid-dependent phosphoglycerate mutase from *Bacillus subtilis*. J. Bacteriol. **182:**4121–4123.
16. **Rigden, D. J., I. Bagyan, E. Lamani, P. Setlow, and M. J. Jedrzejas.** 2001. A cofactor-dependent phosphoglycerate mutase homolog from Bacillus stearothermophilus is actually a broad specificity phosphatase. Protein Sci. **10:**1835–1846.
17. **Sharkey, T. D., S. Yeh, A. E. Wiberley, T. G. Falbel, D. Gong, and D. E. Fernandez.** 2005. Evolution of the isoprene biosynthetic pathway in kudzu. Plant Physiol. **137:**700–712.
18. **Sivy, T. L., M. C. Shirk, and R. Fall.** 2002. Isoprene synthase activity parallels fluctuations of isoprene release during growth of Bacillus subtilis. Biochem. Biophys. Res. Commun. **294:**71–75.
19. **Suckow, J., P. Markiewicz, L. G. Kleina, J. Miller, B. Kisters-Woike, and B. Muller-Hill.** 1996. Genetic studies of the Lac repressor. XV. 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. J. Mol. Biol. **261:**509–523.
20. **Wagner, W. P., D. Helmig, and R. Fall.** 2000. Isoprene biosynthesis in Bacillus subtilis via the methylerythritol phosphate pathway. J. Nat. Prod. **63:**37–40.