Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii*

Tom D. Niehaus[a], Shigeru Okada[b,1], Timothy P. Devarenne[c], David S. Watt[d], Vitaliy Sviripa[d], and Joe Chappell[a,1]

[a]Plant Biology Program, University of Kentucky, Lexington, KY 40546-0312; [b]Department of Aquatic Biosciences, University of Tokyo, Tokyo 113-8657, Japan; [c]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843-2128; and [d]Department of Cellular and Molecular Biochemistry, University of Kentucky, Lexington, KY 40536

Edited by Rodney B. Croteau, Washington State University, Pullman, WA, and approved June 8, 2011 (received for review April 20, 2011)

Botryococcene biosynthesis is thought to resemble that of squalene, a metabolite essential for sterol metabolism in all eukaryotes. Squalene arises from an initial condensation of two molecules of farnesyl diphosphate (FPP) to form presqualene diphosphate (PSPP), which then undergoes a reductive rearrangement to form squalene. In principle, botryococcene could arise from an alternative rearrangement of the presqualene intermediate. Because of these proposed similarities, we predicted that a botryococcene synthase would resemble squalene synthase and hence isolated squalene synthase-like genes from *Botryococcus braunii* race B. While *B. braunii* does harbor at least one typical squalene synthase, none of the other three squalene synthase-like (SSL) genes encodes for botryococcene biosynthesis directly. SSL-1 catalyzes the biosynthesis of PSPP and SSL-2 the biosynthesis of bisfarnesyl ether, while SSL-3 does not appear able to directly utilize FPP as a substrate. However, when combinations of the synthase-like enzymes were mixed together, in vivo and in vitro, robust botryococcene (SSL-1+SSL-3) or squalene biosynthesis (SSL1+SSL-2) was observed. These findings were unexpected because squalene synthase, an ancient and likely progenitor to the other *Botryococcus* triterpene synthases, catalyzes a two-step reaction within a single enzyme unit without intermediate release, yet in *B. braunii*, these activities appear to have separated and evolved interdependently for specialized triterpene oil production greater than 500 MYA. Coexpression of the SSL-1 and SSL-3 genes in different configurations, as independent genes, as gene fusions, or targeted to intracellular membranes, also demonstrate the potential for engineering even greater efficiencies of botryococcene biosynthesis.

algae | biofuels | terpene enzymology

Botryococcus braunii is a colony-forming, freshwater green algae reported to accumulate 30–86% of its dry weight as hydrocarbon oils (1). Three distinct races of *B. braunii* have been described based on the types of hydrocarbons that each accumulates (2). Race A accumulates fatty acid-derived alkadienes and alkatrienes (3), race L accumulates the tetraterpene lycopadiene (4), and race B accumulates triterpenes, predominately botryococcene, squalene, and their methylated derivatives (5). The oils accumulate both in intracellular oil bodies and in association with an extracellular matrix (6), which in race B consists largely of long-chain, cross-linked biopolymers formed in part from acetalization of polymethylsqualene diols (7). Di- and tetra-methylated botryococcenes are generally the most abundant triterpenes accumulating in race B with smaller amounts of tetramethylated-squalene (8) and other structural derivatives of squalene and botryococcene that range from C_{31} to C_{37} accumulating to various levels in different strains and in response to variable culture conditions (9). Other polymethylated derivatives such as diepoxytetramethylsqualene (10), botryolins (11), and brauixanthins (12) have also been reported.

B. braunii race B has received significant attention because it is considered an ancient algal species dating back at least 500 MYA and is one of the few organisms known to have directly contributed to the existing oil and coal shale deposits found on Earth (13–15), accounting for up to 1.4% of the total hydrocarbon content in oil shales (16). Secondly, because the hydrocarbon oils of *B. braunii* race B are readily converted to starting materials for industrial chemical manufacturing and high quality fuels under standard hydrocracking/distillation conditions in yields approaching 97% (Fig. 1*A*) (17), race B has been considered a potential production host for renewable petrochemicals and biofuels. However, the slow growth habit of *B. braunii* poses serious limitations to its suitability as a robust biofuel production system. Capture of the genes coding for this unique oil biosynthetic capacity would therefore provide opportunities to engineer this metabolism into other faster growing and potentially higher yielding organisms (18).

Our approach for identifying the triterpene biosynthetic genes in *B. braunii* has relied in large part on the putative similarities in the biosynthetic mechanisms for squalene and botryococcene (19–21). Squalene biosynthesis has been extensively investigated because it is positioned at a putative branch point in the isoprenoid biosynthetic pathway directing carbon flux to sterol metabolism, and thus represents a potential control point for cholesterol biosynthesis in man (22). Evidence for a two-step reaction mechanism catalyzed by squalene synthase has been described (23) (Fig. 1*B*). The initial reaction step consists of a head-to-head condensation of two farnesyl diphosphate (FPP) molecules to form a stable cyclopropyl intermediate, presqualene diphosphate (PSPP) (24, 25). In the second reaction step, PSPP undergoes a reductive rearrangement in the presence of NADPH to yield squalene possessing a C1-C1′ linkage between the two farnesyl substituents (26, 27) (Fig. 1*B*). Poulter (23) also suggested that botryococcene biosynthesis could occur via an analogous reaction mechanism with the initial reaction proceeding through PSPP, followed by a reductive rearrangement yielding a C3-C1′ linkage between the two farnesyl precursors and possessing an ethyl as well as a methyl group at C3 in the final product.

Extensive investigations of squalene synthase including site-direct mutagenesis (28) and structural elucidation of 3-dimensional structure (29) have focused on five highly conserved domains (domains I–V) thought associated with catalysis (30). Many studies have also utilized these highly conserved domains as a means for isolating the corresponding genes from a diverse range of organisms. For instance, we previously described the functional characterization of a squalene synthase gene from *B. braunii* race B (31). In that work, degenerate oligonucleotide

Author contributions: T.D.N., S.O., D.S.W., and J.C. designed research; T.D.N., S.O., D.S.W., and V.S. performed research; D.S.W. and V.S. contributed new reagents/analytic tools; T.D.N., S.O., T.P.D., D.S.W., and J.C. analyzed data; and T.D.N. and J.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: DNA sequence information for SSL-1, SSL-2, and SSL-3 has been deposited in the GenBank database, www.ncbi.nlm.nih.gov/genbank/ (accession nos. HQ585058, HQ585059, and HQ585060, respectively).

[1]To whom correspondence may be addressed. E-mail: aokada@mail.ecc.u-tokyo.ac.jp or chappell@uky.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1106222108/-/DCSupplemental.

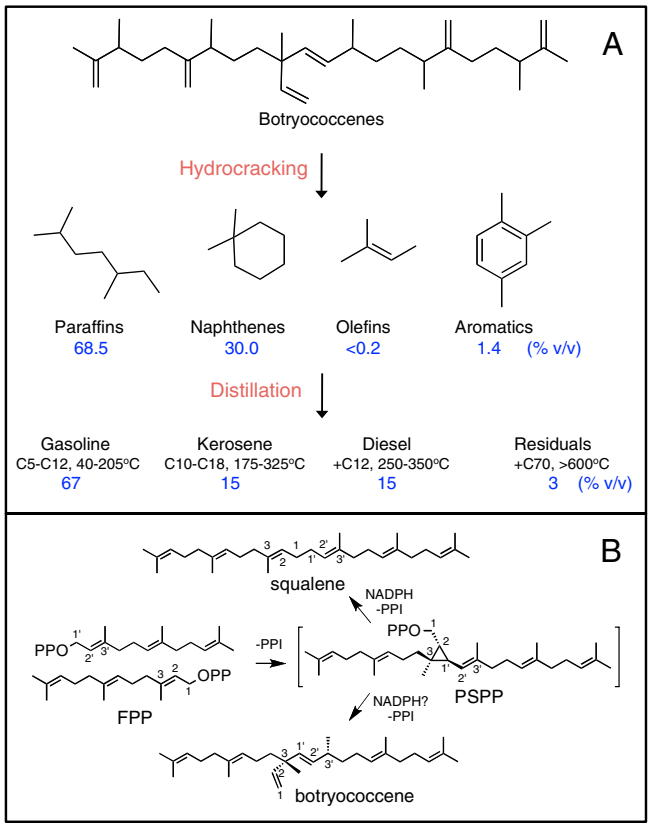

Fig. 1. The triterpene oils of *B. braunii* race B (illustrated as tetramethyl-botryococcene) have been recognized as likely progenitors to existing coal and oil shale deposits for over a century because of geochemical and fossil records (49) and have drawn considerable interest because these oils are readily converted under standard hydrocracking processes to molecular species of direct utility in industrial chemical manufacturing or can be distilled in high yields to all classes of combustible fuels, including gasoline (67%), aviation fuels (15%) and diesel (15%) (carbon chain length, distillation temperature, % volume conversion) (17) (*A*). The biosynthetic origin of the *B. braunii* triterpene oils has remained enigmatic. Poulter (23) suggested that the biosynthesis of the botryococcene scaffold could arise from a mechanism similar to that for squalene, a key intermediate in sterol and cyclized triterpene metabolism (*B*). Squalene biosynthesis occurs from an initial head-to-head condensation of two farnesyl diphosphate molecules (FPP) into the stable intermediate presqualene diphosphate (PSPP), followed by a reductive rearrangement to form squalene catalyzed by a single enzyme without release of the PSPP intermediate (34). Botryococcene biosynthesis is suggested to parallel that of squalene in the first half reaction, differing only in the reductive rearrangement of PSPP to yield the methyl/ethyl branched, 1′-3 linked botryococcene product.

primers complementary to several of the conserved domains were used to amplify a small region of a putative squalene synthase gene, and that gene fragment was then used to isolate a full-length cDNA from a cDNA library. Heterologous expression of that cDNA in bacteria and in vitro characterization of the encoded enzyme validated that the cDNA encoded for a squalene synthase enzyme but lacked any detectable botryococcene synthase activity.

The current results represent our additional efforts to define the botryococcene biosynthetic pathway, to capture the genes coding for these unique enzymological transformations, and to reconstruct the initial steps of these unusual triterpene pathways in a heterologous host.

Results

Functional Identification of Genes for Triterpene Biosynthesis. Because we surmised that a botryococcene synthase enzyme might possess amino acid domains in common with squalene synthase, the *B. braunii* squalene synthase cDNA was used to rescreen the *B. braunii* cDNA library under low stringency hybridization conditions, and a unique squalene synthase-like gene (SSL-1) was isolated and characterized. The SSL-1 gene predicted a squalene synthase-like protein exhibiting some resemblance to other squalene synthase enzymes within domains I–V, but missing a carboxy-terminal, membrane spanning domain (Fig. S1). Surprisingly, purified bacterial-expressed SSL-1 protein did not exhibit either squalene nor botryococcene biosynthesis when assayed in vitro even in the presence of a variety of reducing cofactors like NADPH (Figs. S2*A* and S3), ferredoxin, or cytochrome B5 systems. However, when SSL-1 was expressed in a yeast engineered for high-level production of FPP and having its endogenous squalene synthase and squalene epoxidase genes inactivated, presqualene alcohol (PSOH), the dephosphorylated form of PSPP, accumulated to significant levels (Fig. 2*B*). Subsequent incubations of SSL-1 with radiolabeled FPP confirmed robust in vitro production of PSPP as the sole reaction product with a K_m for FPP of 12.8 μM and catalytic turnover rate (k_{cat}) equal to 2.7×10^{-2}/sec with no stimulation of activity by NADPH addition (Fig. 2*J* and Fig. S3). This suggested that SSL-1 was catalytically competent for the first half reaction of squalene synthase but perhaps required additional conditions or algal factors for complete catalytic activity. Mixing the purified SSL-1 enzyme with algal cell-free lysate did indeed enhance NAD(P)H-dependent botryococcene biosynthesis up to 10-fold, which was also proportional to the amount of the purified SSL-1 protein or the algal lysate added (Fig. S2 *A*–*C*). The mechanism for botryococcene biosynthesis thus appeared to be similar to squalene synthase in its first half reaction, catalysis of PSPP formation, but differed in requiring another algal cofactor that either shuttled reducing equivalents to the reaction mechanism of SSL-1 or participated directly in the conversion of PSPP to botryococcene.

Because no natural occurring squalene synthase catalyzing only the first or second half reactions has been reported, we reasoned that other squalene synthase-like cDNAs for botryococcene biosynthesis might exist and therefore undertook a more exhaustive assessment of the SSL genes expressed in the *Botryococcus braunii* race B cells. The transcriptomic data from two independent sequencing efforts were thus assembled together and screened computationally for additional squalene synthase-like genes. Two additional SSL genes were uncovered and labeled SSL-2 and SSL-3 (Fig. S1). Although both of the predicted proteins showed amino acid sequence similarity to other squalene synthases in excess of 62%, neither bacterial-expressed, purified enzymes exhibited any botryococcene biosynthesis and only SSL-2 showed a low capacity for squalene biosynthesis when incubated with FPP as substrate (Fig. 2*G*). When expressed in yeast, SSL-3 also did not cause the accumulation of any distinct products (Fig. 2*D*), but SSL-2 resulted in the accumulation of a small amount of squalene (approximately 10% of the total) and a terpene compound of unknown structure (Fig. 2*C*). The dominant terpene accumulating in the SSL-2 expressing yeast was subsequently identified by NMR as bisfarnesyl ether and confirmed by comparative analysis of corresponding ether prepared by chemical synthesis (Fig. S4). Subsequent analysis of the reaction products generated by in vitro incubation of SSL-2 with FPP also verified this enzyme as the source of this unique terpene ether (Fig. 2 *G* and *J*).

The observations of unique terpene products from squalene synthase-like enzymes in *Botryococcus*, namely PSPP by SSL-1 and bisfarnesyl ether by SSL-2, suggested that triterpene metabolism in this algae may operate differently from that in other organisms. Hence, we considered the possibility that multiple SSL proteins might be required to give botryococcene biosynthesis. To evaluate this possibility, the different SSL genes were coexpressed in yeast, or the heterologous expressed and purified proteins were incubated in various combinations. When SSL-1 was coexpressed with SSL-2, the amount of squalene accumulat-

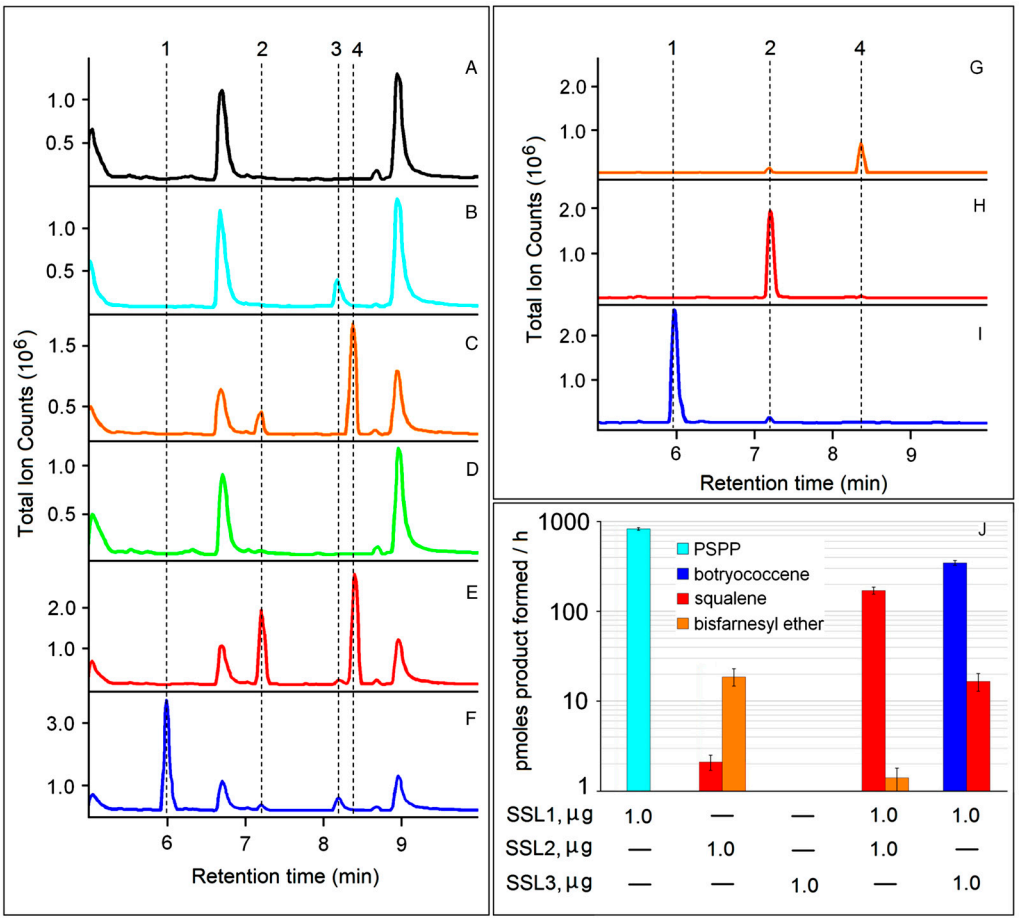

Fig. 2. Functional characterization of the squalene synthase-like genes of *Botryococcus braunii* race B. The squalene synthase-like genes, SSL-1, SSL-2 and SSL-3, were expressed in yeast separately [SSL-1 (*B*), SSL-2 (*C*), or SSL-3 (*D*)] or in combinations [SSL-1 + SSL-2 (*E*), SSL-1 + SSL-3 (*F*)] and the hexane extractable metabolites profiled by GC-MS. The chemical profile of yeast not engineered with any gene constructs serves as the background control (*A*). The SSL genes were also expressed in bacteria, the affinity-tagged proteins purified and assayed separately [SSL-2 (*G*)] or in combinations [SSL-1 + SSL-2 (*H*); SSL-1 + SSL-3 (*I*)] for the reaction products generated upon incubation with FPP and profiled by GC-MS (*G–I*), or for quantitative determination of radiolabeled FPP incorporated into specific reaction products separated by TLC (*J*). Data (*J*) represents mean ± S.E.M. obtained from three independent experiments ($n = 3$). The chromatograms (*A–I*) are also annotated for the elution behavior of botryococcene (1), squalene (2), presqualene alcohol (3), and bisfarnesyl ether (4).

ing increased about 30-fold (Fig. 2*E*) along with a significant accumulation of bisfarnesyl ether still occurring. When purified SSL-1 and SSL-2 enzymes were incubated in a 1:1 stoichiometric ratio, squalene accumulation predominated (Fig. 2*H*), suggesting that something different mechanistically might be occurring when the SSL-1 and 2 genes were coexpressed in yeast (see below). More surprising, however, when SSL-1 and SSL-3 were coexpressed, botryococcene accumulation became readily apparent and accumulated to upwards of 20 mg/L along with 0.7 mg/L of squalene (Fig. 2*F*). In vitro incubations of the purified SSL-1 and SSL-3 proteins confirmed botryococcene as the predominant reaction product with squalene representing only 3–4% of the total reaction products (Fig. 2 *I* and *J*). Additional in vitro studies have also confirmed that both SSL-2 and SSL-3 are able to efficiently catalyze the biosynthesis of squalene and botryococcene, respectively, from PSPP but not FPP, and these activities of SSL-2 and SSL-3 are sufficient to account for all the squalene and botryococcene biosynthesized in combined assays with SSL-1 (Fig. S5).

Mechanistic Considerations for Bisfarnesyl Ether Biosynthesis. When incubated by itself, SSL-2 catalyzes the NADPH-dependent biosynthesis of approximately 90% bisfarnesyl ether and 10% squalene (Fig. 2 *G* and *J*). This suggests that SSL-2 does have the ability to generate PSPP but at a much lower efficiency relative to ether formation. Based on a consideration of the detailed carbocation mechanism elucidated for the biosynthesis of squalene from FPP (Fig. 1*B*) (23), one might not expect bisfarnesyl ether biosynthesis to involve a PSPP intermediate. Instead, if the initial carbocation generated on one of the two SSL-2 bound FPP molecules were quenched by reaction with an available water molecule, and if the so formed farnesol (FOH) were positioned in the correct orientation and proximity to a second FPP molecule, displacement of the pyrophosphate group via a S_N2 Williamson ether synthesis-type reaction (32) could yield the bisfarnesyl ether (Fig. 3). Support for such a mechanism comes from the incorporation of radiolabeled FOH directly into the bisfarnesyl ether product, but only when SSL-2 is incubated with both FOH and FPP (Table 1).

To determine if the mechanism of NADPH dependence for bisfarnesyl ether formation by SSL-2 was catalytic or structural, the quantitative yield of reaction product and NADPH oxidation were determined. While the biosynthesis of 1,072 pmoles of squalene was correlated with an equal stoichiometric oxidation of 1,098 pmoles of NADPH by the *Nicotiana benthamiana* squalene synthase enzyme, greater than 21 pmoles of bisfarnesyl ether were formed by SSL-2 when only 4.6 pmoles of NADPH were oxidized (Table S1). Approximately half of the NADPH oxidation by SSL-2 under these conditions could be associated with the biosynthesis of 2.2 pmoles of squalene (Table S1). Hence, about 10 times more bisfarnesyl ether is formed per mole equivalent

Table 1. Substrate specificity of SSL-2

Substrate	Bisfarnesyl ether formed (pmoles/h•µg protein)
^{3}H – farnesol	0
FPP + ^{3}H – farnesol	16.3 +/– 0.4
^{3}H – FPP	13.5 +/– 3.3

Purified SSL-2 protein (2 µg) was incubated with either 10 µM 1 – ^{3}H – farnesol, 10 µM 1 – ^{3}H – farnesol plus 10 µM FPP, or 10 µM 1-3H-FPP in a 50-µl reaction at 37 °C for 1 h, the reaction products extracted with MTBE, and aliquots separated on silica TLC plates. The radioactivity incorporated in the zones corresponding to bisfarnesyl ether were determined by scintillation counting. Data represents mean ± SEM.

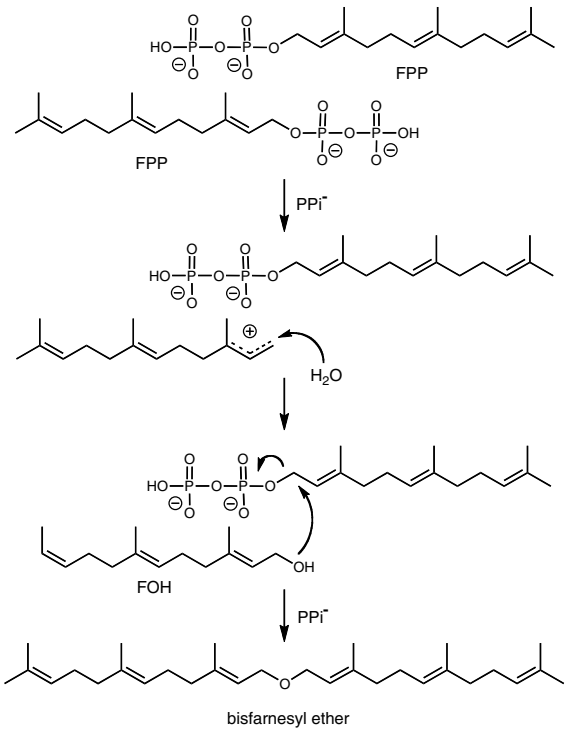

Fig. 3. Proposed mechanism for bisfarnesyl ether biosynthesis by SSL-2. When two molecules of FPP are bound by the SSL-2 enzyme, ionization of the diphosphate substituent from one creates a carbocation, which can react with a water molecule in close proximity to generate farnesol, FOH. If the FOH becomes appropriately positioned relative to the second FPP molecule, then a Williamson ether synthesis (32) reaction could occur to yield bisfarnesyl ether.

of NADPH oxidation (21 pmoles versus 2.4 pmoles), more consistent with an allosteric or structural role for NADPH in the SSL-2 bisfarnesyl ether reaction rather than a catalytic one. A similar role for NADPH in stimulating PSPP formation by squalene synthases was reported earlier (29, 33, 34).

Improving the Efficiency of Botryococcene Biosynthesis. Production of botryococcene by yeast was improved by engineering different configurations of the SSL-1 and SSL-3 genes (Fig. 4). While coexpression of SSL-1 and SSL-3 yielded significant botryococcene, peptide fusions of SSL-1 and SSL-3 connected by a triplet repeat linker of GGSG improved production capacity greater than twofold to upwards of 50 mg/L. Further enhancement to over 70 mg/L was observed by appending the carboxy-terminal 63 or 71 amino acids of the *Botryococcus* squalene synthase onto the carboxy-termini of SSL-1 and SSL-3 enzymes, respectively. These terminal amino acids serve to tether squalene synthase, and by inference SSL-1 and 3, to the yeast's endo-membrane system, which might bring the enzymes in closer proximity to one another or give the enzymes greater access to endogenous FPP pools. Further support for this notion has been the observation of greater than 100 mg/L of botryococcene by yeast overexpressing gene fusions of SSL-1 and SSL-3 harboring the putative ER membrane targeting sequence of the botryococcus squalene synthase.

Discussion

The results presented here were unexpected because squalene biosynthesis is known as a two-step process catalyzed by a single enzyme (Fig. 5). FPP is first converted to the intermediate PSPP, followed by its reductive rearrangement to squalene (24). However, PSPP is not evident in these reactions unless NADPH, the reducing reagent, is omitted from the incubations (19). Under

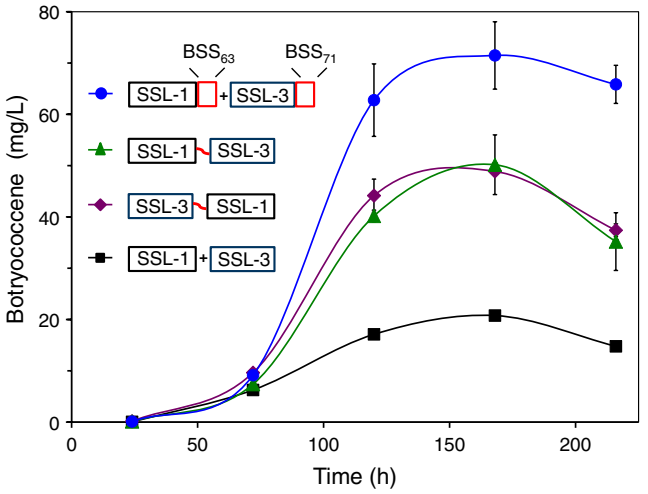

Fig. 4. Comparison of botryococcene production in yeast engineered with different configurations of SSL-1 and SSL-3. Yeast line TN7 was engineered with the SSL-1 and SSL-3 genes on separate plasmids (squares), with gene fusions [SSL-1 fused to SSL-3 via a triplet repeat of GGSG (triangles), or vice versa (diamonds)], or with 63 or 71 amino acids of the carboxy terminus of the *Botryococcus* squalene synthase, sequences containing a membrane spanning domain, appended to the carboxy termini of the SSL-1 and SSL-3 enzymes, respectively (circles). The data represents mean ± S.E.M.

conditions of adequate NADPH, it appears unlikely that PSPP is released from the squalene synthase enzyme, then rebound as a natural consequence of the catalytic cycle (34). Regardless, a single enzyme is responsible for the entire conversion process and this mechanism appears highly conserved from yeast to man, including algae like *Botryococcus* (31). In contrast, botryococcene biosynthesis appears to require the successive action of two distinct enzymes. First SSL-1 catalyzes the biosynthesis of PSPP as a separate and distinct product, which the second enzyme, SSL-3, efficiently converts to botryococcene in a NADPH-dependent manner. Whatever the evolutionary forces driving this division of labor might have been, it also appears to have occurred twice within the life history of *Botryococcus*. When SSL-1 is coexpressed with SSL-2, squalene accumulates, which we speculate might represent a distinct pool of squalene in *Botryococcus* destined to specialized roles like the biogenesis of the extracellular matrix and other squalene derivatives.

Support for the neofunctionalization of these unusual binary systems for triterpene biosynthesis is provided by the distinctive biosynthetic activities associated with SSL-2 (Fig. 5). First, this enzyme catalyzes the NADPH-dependent biosynthesis of an unusual terpene ether. There are no reports of bisfarnesyl ether accumulation in *Botryococcus* or any other organism, but it could be incorporated into other more complex matrix polymers masking its detection. One possible means for bisfarnesyl ether biosynthesis does not involve a PSPP intermediate but instead an alternative reactivity of two bound farnesyl moieties via a S_N2 Williamson ether synthesis-type reaction (Fig. 3) (32). Support for such a mechanism comes from the incorporation of radiolabeled FOH directly into the bisfarnesyl ether product, but only when SSL-2 is incubated with both FOH and FPP (Table 1). Second, the accumulation of both squalene and bisfarnesyl ether in yeast coexpressing SSL-1 and SSL-2 is also consistent with this proposed mechanism. The yeast line used for these studies is engineered for high FPP production but tends to accumulate FOH as a consequence of FPP dephosphorylation catalyzed by endogenous phosphatases (35, 36). Hence, the yeast coexpressing SSL-1 and SSL-2 have significant pools of FOH and FPP, which will compete with any PSPP generated by SSL-1 for binding and catalysis by SSL-2. Third, while there is no obvious or direct chemical requirement for reducing equivalents in the biosynthesis of

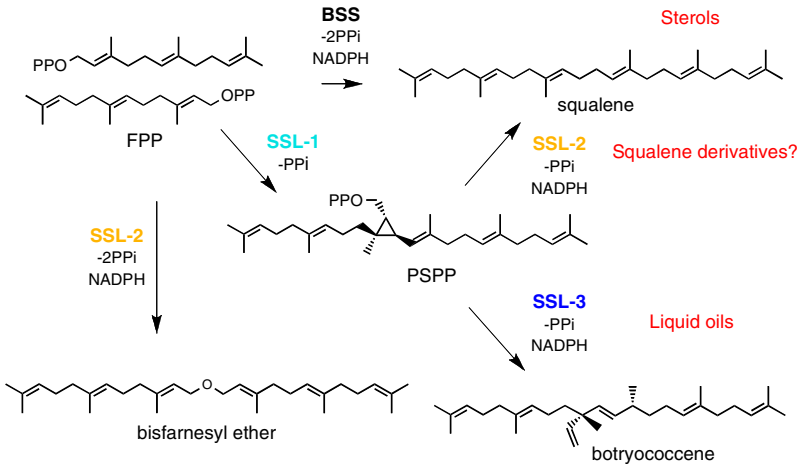

Fig. 5. A cartoon depiction of the catalytic roles of the squalene synthase-like enzymes in *Botryococcus braunii* race B and their putative contributions to the triterpene constituents that accumulate. The previously identified squalene synthase gene (BSS) (31) is thought to provide squalene essential for sterol metabolism, whereas the squalene synthase-like genes SSL-1, SSL-2, and SSL-3 provide for the triterpene oils serving specialized functions for the algae. In combination with SSL-1, SSL-2 could provide squalene for extracellular matrix and methylated squalene derivatives, while SSL-1 plus SSL-3 generates botryococcene, which along with its methyl derivatives, accounts for the majority of the triterpene oil.

the bisfarnesyl ether from FPP and FOH, the significance of the NADPH dependence might relate to a structural role rather than a catalytic one. Pandit et al. (29) suggested that NADPH binding to its putative bind site in the human squalene synthase might stabilize a region of the enzyme not well resolved in the crystal structure, and thus positioning a domain into close association with the active site. NADPH binding to the SSL-2 enzyme could evoke a similar conformational change that renders the SSL-2 enzyme competent for either bisfarnesyl ether or squalene biosynthesis dependent on available substrates (FPP, FOH, and PSPP). Hence, not only has SSL-2 maintained its catalytic ability to convert PSPP to squalene, it has evolved a novel catalytic activity yielding a bisprenyl ether from prenyl diphosphates.

One possibility for how these unique triterpene synthases arose is that a progenitor squalene synthase gene could have duplicated to yield multiple gene copies. While one copy (BSS) maintained its coding capacity for squalene synthase activity, essential for sterol metabolism, the other copies (SSL-1, SSL-2, and SSL-3) would have afforded opportunities for evolutionary diversification. Alternatively, *Botryococcus* could have acquired multiple copies of SSL genes by a horizontal gene transfer process and those genes may have evolved specialized synthase-like activities. For example, one of the acquired squalene synthase-like genes could have evolved the capacity for botryococcene biosynthesis and a subsequent gene duplication event could have resulted in loss of function for either the first half reaction or the second. No matter the specific mechanism, what makes the possible events associated with the neofunctionalization of the SSL enzymes particularly intriguing is that specialized triterpene oil accumulation, like botryococcene, could not have occurred without both SSL-1 and SSL-3 evolving in concert with one another.

There are other examples of similar division and diversification of enzymological capacities within key genes for pyrimidine (37), diterpene (38), and triterpene (39) metabolism. For instance, biosynthesis of the diterpene kaurene in many fungi relies on a single, multifunctional enzyme (40) that catalyzes the conversion of the linear isoprenoid intermediate geranylgeranyl diphosphate to the bicyclic copalyl diphosphate (CPP) product. CPP then undergoes a second cyclization reaction initiated at a separate binding site on the same enzyme to yield kaurene. In higher plants, the enzymes for CPP and kaurene biosynthesis are encoded by separate and distinct genes (38). Specific CPP synthases within rice catalyze the biosynthesis of either *ent*-CPP or *syn*-CPP isomers (41, 42). These are complemented with equally distinct diterpene synthases that can utilize one or the other CPP isomer for hormone or defense compound biosynthesis (43, 44). Yet, there are other diterpene synthases that have retained these two enzyme functions but have evolved whole new catalytic outcomes (45). Osbourn and coworkers (39, 46) have also provided evidence that the genes encoding for the enzymes catalyzing the cyclization of oxidosqualene to distinct tetra- and penta-cyclic classes of triterpenes, primarily sterols and defense related saponins, respectively, likely arose from common ancestor genes evolving novel catalytic functions dedicated to primary and specialized metabolism. Microbial forms of dihydrosqualene synthase, like CrtM, might also be considered an example of squalene synthase-like enzyme diversification (47, 48). CrtM relies on PSPP biosynthesis but does not utilize NADPH for the second half reaction. CrtM instead yields dehydrosqualene, a reaction product with much in common with phytoene, the tetraterpene equivalent of dehydrosqualene, and by inference shares catalytic features of the second half reaction in common with phytoene synthase. Nonetheless, what distinguishes the current results from all the others is there are no other known examples where the half-reaction specificity of squalene synthases appear separated from one another and subject to evolutionary diversification, except for that reported here for *Botryococcus*.

The family of squalene synthase-like enzymes in *Botryococcus* is also informative relative to the recent elucidation of the crystal structure of dehydrosqualene synthase (CrtM) of *Staphylococcus aureus*, a target enzyme for a new generation of antiinfective reagents, along with refinements in the human squalene synthase structure (47, 48). Those studies detailed how two FPP molecules bind to CrtM and human squalene synthase, are converted to the PSPP intermediate, and then repositioned in the active site pocket in preparation for the second half reaction. Key residues identified include those that coordinate magnesium ions for their interactions with the diphosphate substituents of the FPPs and PSPP, and hence considered involved in both half-reactions. Based on sequence alignments (Fig. S1), many of these residues (S19, Y41, R45, D48, D52, Y129, N168, and D177, numbering according to CrtM and annotated by a star above the residue in Fig. S1) appear conserved in the *Botryococcus* squalene synthase and all three of the SSL enzymes. Because SSL-2 and SSL-3 are deficient in PSPP biosynthesis, these particular residues are not by themselves sufficient for PSPP biosynthesis. Conversely, since SSL-1 can only catalyze the formation of PSPP, these same residues do not appear sufficient to initiate the second half reaction. Amino acids at other positions are undoubtedly important for PSPP formation and the catalytic specificity of the second half reaction, squalene versus botryococcene biosynthesis. Experiments to functionally define which amino acids at which positions are responsible for the enzymological specificity of these triterpene synthases will be significantly advantaged by having these unique *Botryococcus* SSL enzymes, which are specialized to either the first half reaction or the second.

Altogether, our results establish that botryococcene and squalene oils are synthesized in *Botryococcus braunii* race B by

the combined action of separate and distinct squalene synthase-like enzymes, have opened up new avenues for understanding the chemical specificity and diversification within this class of enzymes and provide a demonstration for the bioengineering and production of a key petrochemical replacement.

Methods

The squalene synthase-like cDNAs were isolated either by screening a *Botryococcus* cDNA library using low stringency hybridization conditions with a radiolabeled *Botryococcus* squalene synthase probe (yielding SSL-1), or by computational screening of the combined *Botryococcus* transcriptomic datasets with the *Botryococcus* squalene synthase cDNA sequence (yielding SSL-2 and SSL-3). These three genes were inserted into the pET28a vector for bacterial expression and the YEp352 or pESC vectors for yeast expression. Bacterial-expressed enzymes were purified, incubated with FPP or [1-^{3}H] FPP, and hexane extracts analyzed either by GC-MS, or by scintillation counting of the indicated products isolated by TLC, respectively. Various combinations of the SSL genes were transformed into the TN7 yeast line, the transformants grown in either YPDE or SCE media, and organic extracts of the cultures analyzed by GC-MS. TN7 was created by insertional mutagenesis of the ERG1 gene in the Cali-7 yeast line. The unknown terpene accumulating in TN7 expressing SSL-2 was purified by silica-HPLC, then subjected to standard NMR analyses along with chemically synthesized bisfarnesyl ether. Full details are given in *SI Methods*.

ACKNOWLEDGMENTS. We thank Scott Kinison for technical assistance with all facets of the work described here. The CALI-7 yeast line was kindly provided by L.S. Song and R. Rosen of BioTechnical Resources, Manitowoc, WI. This work was supported by grants from the National Science Foundation (CBET-0828817) and Sapphire Energy to J.C.; the National Institutes of Health (NIH) [2P20 RR020171 from the National Center for Research Resources (NCRR)] to D.S.W. (the contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the NIH or the NCRR); and by Grant-in-Aid for Young Scientists (A) (12760137), Grant-in-Aid for Scientific Research (C) (16580166), Grant-in-Aid for Scientific Research (C) (18580202), and Grant-in-Aid for Scientific Research (B) (21380130) from the Japanese Society for the Promotion of Science to S.O.

1. Brown AC, Knights BA, Conway E (1969) Hydrocarbon content and its relationship to physiological state in green alga botryococcus braunii. *Phytochemistry* 8:543–547.
2. Metzger P, Largeau C (2005) Botryococcus braunii: A rich source for hydrocarbons and related ether lipids. *Appl Microbiol Biotechnol* 66:486–496.
3. Gelpi E, Oro J, Schneide HJ, Bennett EO (1968) Olefins of high molecular weight in 2 microscopic algae. *Science* 161:700–701.
4. Metzger P, Allard B, Casadevall E, Berkaloff C, Coute A (1990) Structure and chemistry of a new chemical race of *Botryococcus braunii* (Chlorophyceae) that produces lycopadiene, a tetraterpenoid hydrocarbon. *J Phycol* 26:258–266.
5. Okada S, Murakami M, Yamaguchi K (1995) Hydrocarbon composition of newly isolated strains of the green microalga *Botryococcus braunii. J Appl Phycol* 7:555–559.
6. Weiss TL, et al. (2010) Raman spectroscopy analysis of botryococcene hydrocarbons from the green microalga *Botryococcus braunii. J Biol Chem* 285:32458–32466.
7. Metzger P, Rager MN, Largeau C (2007) Polyacetals based on polymethylsqualene diols, precursors of algaenan in botryococcus braunii race b. *Org Geochem* 38:566–581.
8. Huang Z, Poulter CD (1989) Tetramethylsqualene, a triterpene from *Botryococcus braunii* var showa. *Phytochemistry* 28:1467–1470.
9. Metzger P, Berkaloff C, Casadevall E, Coute A (1985) Alkadiene-producing and botryococcene-producing races of wild strains of *Botryococcus braunii. Phytochemistry* 24:2305–2312.
10. Metzger P (1999) Two terpenoid diepoxides from the green microalga *Botryococcus braunii*: Their biomimetic conversion to tetrahydrofurans and tetrahydropyrans. *Tetrahedron* 55:167–176.
11. Metzger P, Rager MN, Largeau C (2002) Botryolins A and B, two tetramethylsqualene triethers from the green microalga *Botryococcus braunii. Phytochemistry* 59:839–843.
12. Okada S, Tonegawa I, Matsuda H, Murakami M, Yamaguchi K (1997) Braunixanthins 1 and 2, new carotenoids from the green microalga *Botryococcus braunii. Tetrahedron* 53:11307–11316.
13. Derenne S, et al. (1997) Chemical structure of the organic matter in a pliocene maar-type shale: Implicated *Botryococcus* race strains and formation pathways. *Geochim Cosmochim Acta* 61:1879–1889.
14. Glikson M, Lindsay K, Saxby J (1989) *Botryococcus*—a planktonic green-alga, the source of petroleum through the ages—transmission electron microscopical studies of oil shales and petroleum source rocks. *Org Geochem* 14:595–608.
15. Mastalerz M, Hower JC (1996) Elemental composition and molecular structure of *Botryococcus* alginite in westphalian cannel coals from Kentucky. *Org Geochem* 24:301–308.
16. Moldowan JM, Seifert WK (1980) 1st discovery of botryococcane in petroleum. *J Chem Soc Chem Comm* 19:912–914.
17. Hillen LW, Pollard G, Wake LV, White N (1982) Hydrocracking of the oils of *Botryococcus braunii* to transport fuels. *Biotechnol Bioeng* 24:193–205.
18. Banerjee A, Sharma R, Chisti Y, Banerjee UC (2002) *Botryococcus braunii*: A renewable source of hydrocarbons and other chemicals. *Crit Rev Biotechnol* 22:245–279.
19. Jarstfer MB, Zhang DL, Poulter CD (2002) Recombinant squalene synthase. Synthesis of non-head-to-tail isoprenoids in the absence of NADPH. *J Am Chem Soc* 124:8834–8845.
20. Pan JJ, Bugni TS, Poulter CD (2009) Recombinant squalene synthase. Synthesis of cyclopentyl non-head-to-tail triterpenes. *J Org Chem* 74:7562–7565.
21. Zhang DL, Poulter CD (1995) Biosynthesis of non-head-to-tail isoprenoids—synthesis of 1′-1-structures and 1′-3-structures by recombinant yeast squalene synthase. *J Am Chem Soc* 117:1641–1642.
22. Bergstrom JD, et al. (1993) Zaragozic acids—a family of fungal metabolites that are picomolar competitive inhibitors of squalene synthase. *Proc Natl Acad Sci USA* 90:80–84.
23. Poulter CD (1990) Biosynthesis of non-head-to-tail terpenes—formation of 1′-1 and 1′-3 linkages. *Acc Chem Res* 23:70–77.
24. Rilling HC (1966) A new intermediate in biosynthesis of squalene. *J Biol Chem* 241:3233–3236.
25. Sasiak K, Rilling HC (1988) Purification to homogeneity and some properties of squalene synthetase. *Arch Biochem Biophys* 260:622–627.
26. Blagg BSJ, Jarstfer MB, Rogers DH, Poulter CD (2002) Recombinant squalene synthase. A mechanism for the rearrangement of presqualene diphosphate to squalene. *J Am Chem Soc* 124:8846–8853.
27. Jarstfer MB, Blagg BSJ, Rogers DH, Poulter CD (1996) Biosynthesis of squalene. Evidence for a tertiary cyclopropylcarbinyl cationic intermediate in the rearrangement of presqualene diphosphate to squalene. *J Am Chem Soc* 118:13089–13090.
28. Gu PD, Ishii Y, Spencer TA, Shechter I (1998) Function-structure studies and identification of three enzyme domains involved in the catalytic activity in rat hepatic squalene synthase. *J Biol Chem* 273:12515–12525.
29. Pandit J, et al. (2000) Crystal structure of human squalene synthase—a key enzyme in cholesterol biosynthesis. *J Biol Chem* 275:30610–30617.
30. Robinson GW, Tsay YH, Kienzle BK, Smithmonroy CA, Bishop RW (1993) Conservation between human and fungal squalene synthetases—similarities in structure, function, and regulation. *Mol Cell Biol* 13:2706–2717.
31. Okada S, Devarenne TP, Chappell J (2000) Molecular characterization of squalene synthase from the green microalga *Botryococcus braunii*, race B. *Arch Biochem Biophys* 373:307–317.
32. Houben J, Weyl T, Müller E (1965) *Methoden der organischen chemie* (G. Thieme, Stutgart, Germany), 4th Ed, p 832.
33. Agnew WS, Popjak G (1978) Squalene synthetase—stoichiometry and kinetics of presqualene pyrophosphate and squalene synthesis by yeast microsomes. *J Biol Chem* 253:4566–4573.
34. Radisky ES, Poulter CD (2000) Squalene synthase: Steady-state, pre-steady-state, and isotope-trapping studies. *Biochemistry* 39:1748–1760.
35. Song LS (2003) Detection of farnesyl diphosphate accumulation in yeast erg9 mutants. *Anal Biochem* 317:180–185.
36. Takahashi S, et al. (2007) Metabolic engineering of sesquiterpene metabolism in yeast. *Biotechnol Bioeng* 97:170–181.
37. Nara T, Hshimoto T, Aoki T (2000) Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. *Gene* 257:209–222.
38. Peters RJ (2010) Two rings in them all: The labdane-related diterpenoids. *Nat Prod Rep* 27:1521–1530.
39. Field B, Osbourn AE (2008) Metabolic diversification—independent assembly of operon-like gene clusters in different plants. *Science* 320:543–547.
40. Toyomasu T, et al. (2000) Cloning of a full-length cDNA encoding ent-kaurene synthase from *Gibberella fujikuroi*: Functional analysis of a bifunctional diterpene cyclase. *Biosci Biotechnol Biochem* 64:660–664.
41. Prisic S, Xu MM, Wilderman PR, Peters RJ (2004) Rice contains two disparate ent-copalyl diphosphate synthases with distinct metabolic functions. *Plant Physiol* 136:4228–4236.
42. Xu MM, Hillwig ML, Prisic S, Coates RM, Peters RJ (2004) Functional identification of rice syn-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. *Plant J* 39:309–318.
43. Xu MM, Wilderman PR, Peters RJ (2007) Following evolution's lead to a single residue switch for diterpene synthase product outcome. *Proc Natl Acad Sci USA* 104:7397–7401.
44. Xu MM, et al. (2007) Functional characterization of the rice kaurene synthase-like gene family. *Phytochemistry* 68:312–326.
45. Cao R, et al. (2010) Diterpene cyclases and the nature of the isoprene fold. *Proteins* 78:2417–2432.
46. Haralampidis K, et al. (2001) A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. *Proc Natl Acad Sci USA* 98:13431–13436.
47. Lin F-Y, et al. (2010) Mechanism of action and inhibition of dehydrosqualene sythase. *Proc Natl Acad Sci USA* 107:21337–21342.
48. Liu CI, et al. (2008) A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus virulence. Science* 319:1391–1394.
49. Traverse A (1955) Occurrence of the oil-forming alga *Botryococcus* in lignites and other tertiary sediments. *Micropaleontology* 1:343–350.