Journal of Petroleum & Environmental Biotechnology

ISSN: 2157-7463

Elumalai et al., J Pet Environ Biotechnol 2018, 9:3
DOI: 10.4172/2157-7463.1000371

Research Article Open Access

In Silico Modeling and Characterization of Squalene Synthase and Botryococcene Synthase Enzymes from a Green Photosynthetic Microalga *Botryococcus braunii*

Elumalai S[1*], Sangeetha T[1] and Rajesh Kanna G[2]

[1]*Department of Biotechnology, University of Madras, Guindy Campus, Chennai, India*
[2]*Department of Plant Biology and Plant Biotechnology, Presidency College (Autonomous), Chennai, India*

Abstract

The green photosynthetic microalgae are considered as a major source of lipids from lacustrine and marine environments. Among them, *Botryococcus braunii* plays a key role due to its high efficiency production of huge amount of unsaturated hydrocarbons up to 75% of its dry weight. Evidently, more number of new compounds has been reported in sediments as deposits in both the marine and lacustrine environments. These depositions are reported as sediments from the algal lipids during the course of evolution. Botryococcene is one among the long chain hydrocarbon reported in higher amount extracellular as depositions from this microalga *Botryococcus braunii* race B. However, mass cultivation of this microalga for botryococcene as sustainable and renewable biofuel is a challenging target due to its doubling time and slow growth. Therefore, genetic engineering may play a key role to solve this issue. In addition to squalene synthase, squalene synthase-like genes have been reported from the race B of *B. braunii* which are SSL-1 (Presqualene diphosphate synthase), SSL-2 (Botryococcus squalene synthase) and SSL-3 (Botryococcene synthase) genes. This is an astounding report that these genes are controlling the production of long chain hydrocarbon botryococcene. Since, our present study clearly reveals that the squalene synthase and botryococcene synthase of *B. braunii* BB1 strain have very low protein homology of below 50% with human squalene synthase. Thus, it is clear that no high resolution studies have been conducted yet on these important enzymes. Even though, many overexpression studies have been carried out on these enzymes, x-ray diffraction studies may yield more information on the enzymes about its enzyme substrate specificity and it may help to improve the stability and efficiency of the enzymes for industrial aspects.

Keywords: *Botryococcus braunii*; Squalene synthase; botryococcene synthase; In silico modeling; characterization

Introduction

The green photosynthetic microalga *Botryococcus braunii* is extol for its three different kinds of hydrocarbon syntheses such as cis and trans C25-C31 n-alkane and C29 triene, a series of C34H58 botryococcene and C40H78 trans, translycopadiene. Similarly, three different races were distinguished by the production of their own unique hydrocarbons respectively as A (Alkadine), B (Botryococcene) and L (Lycopadiene) [1]. These hydrocarbons are found as depositions in sediments and reported in many regions of the world such as El Junco, Galapagos [1], and crude oil shales of Sumatra [2], Coastal bitumen in Australis [3]. Comparatively, a different strain Botryococcus mahabali have been reported for such kind of hydrocarbon depositions in East Coast of India [4]. Therefore, in recent decades more number of research activities has been carried out to bring up large scale hydrocarbon production from this unique algal species to solve fuel crisis. Since, it has not been achieved due to very slow growth of the alga and various other factors.

Niehaus et al. in the year 2011 [5] reported three different squalene synthases like enzymes such as Presqualene diphosphate synthase (SSL-1), Botryococcus squalene synthase (SSL-2) and Botryococcene synthase (SSL-3) in addition to squalene synthase enzyme which altogether controls the biological production of squalene and botryococcene in *Botryococcus braunii* (Figure 1). Among those enzymes, presqualene diphosphate synthase (SSL-1) and botryococcus squalene synthase (SSL-2) are controlling the biosynthesis of squalene and the rest one botryococcene synthase (SSL-3) is responsible for the production of botryococcene.

The squalene synthase (EC: 2.5.1.21) performs a dual role by condensing two molecules of farnesyl diphosphate (FPPs) to form Pre-Squalene Diphosphate (PSPP) and then further convert it to synthesize squalene by utilizing NADPH and Mg^{2+} [6]. Similarly, presqualene diphosphate synthase (SSL-1) (2.5.1.103) also condenses two molecules of FPPs to synthesize PSPP and botryococcus squalene synthase (SSL-2) (1.3.1.96) to synthesize squalene and one more key enzyme is botryococcene synthase (SSL-3) (1.3.1.97) expressed to synthesize C30 botryococcene. The botryococcene is the precursor for the hydrocarbon which can be hydrolyzed and used as fuel for combustion engines [7].

There are some variations in the mechanism of both the enzymes responsible for the biosynthesis of squalene and botryococcene. Squalene can be synthesized by the enzyme squalene synthase which involves two step reactions; first it condenses two molecules of farnesyl diphosphate to form presqualene diphosphate [8,9] and in the second step it cleaves the cyclopropane ring of presqualene diphosphate to form 10-1 linkage by the reduction of NADPH. Similarly, C30 botryococcene also synthesized by two step reaction, in which the second step; the cleavage of cyclopropane ring of presqualene diphosphate alone differs to form 10-3 linkage.

Some studies on the synthesis of squalene reported that the

*Corresponding author: Elumalai S, Department of Biotechnology, University of Madras, Guindy Campus, Chennai, India, Tel: + 91 2539 9422; E-mail: ananandal@gmail.com

Received May 16, 2018; **Accepted** July 07, 2018; **Published** July 25, 2018

Citation: Elumalai S, Sangeetha T, Rajesh Kanna G (2018) In Silico Modeling and Characterization of Squalene Synthase and Botryococcene Synthase Enzymes from a Green Photosynthetic Microalga *Botryococcus braunii*. J Pet Environ Biotechnol 9: 371. doi: 10.4172/2157-7463.1000371

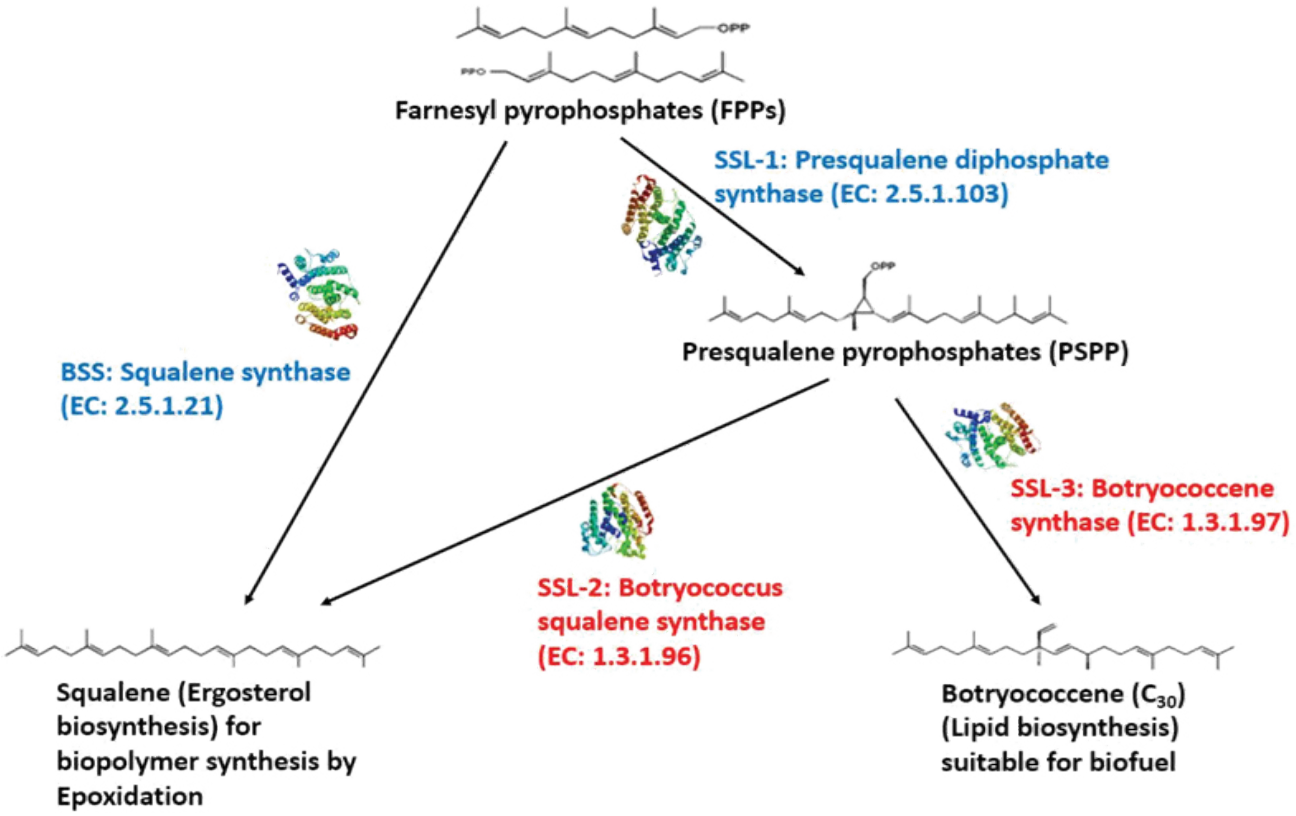


Figure 1: The metabolic pathway depicting the role of different enzymes, control the biosynthesis of squalene and Botryococcene; in which, BSS and SSL-1 are classified under transferase and SSL-2 and SSL-3 belongs to oxidoreductase based on the enzyme classification [5].

recombinant squalene synthase enzyme from yeast was incubated in the absence of NADPH have synthesized 10-3 linked (10S, 13S)-10-hydroxybotryococcene in addition to the two 10-1 linked squalene derivatives (Z)-dehydrosqualene and (R)-12-hydroxysqualene [10]. This resulted that the squalene synthase enzyme under various conditions can synthesize botryococcene. At the same time, there is a separate enzyme found in the micro algae *Botryococcus braunii* for the C30 botryococcene biosynthesis.

The SSL-1 catalyzes the first step of the reaction carried out by squalene synthase by converting farnesyl diphosphate into presqualene diphosphate; the SSL-2 converts the presqualene diphosphate into squalene and the SSL-3 convert the presqualene diphosphate into C30 botryococcene. At this stage, the SSL-2 and SSL-3 are the key enzymes in order to determine the synthesis of squalene or C30 botryococcene biosynthesis respectively (Figure 1). The C30 botryococcene can be further methylated from C31 to C37 botryococcenes [5] to extract free hydrocarbons as biofuels. The major objectives involves the Squalene synthase gene sequencing, amino acid sequence analysis, phylogenetic studies and superimposed molecular structural studies on the four genes squalene synthase, SSL-1, SSL-2 and SSL-3 from three different strains of the micro alga *Botryococcus braunii*.

Materials and Methods

Isolation of RNA

Before isolation of RNA from the microalga, the pure culture of microalga *Botryococcus braunii* BB1 strain was maintained *In vitro*. From the pure culture, about 0.5 ml of well-grown pure culture was ground well into fine powder with liquid nitrogen in a mortar and pestle. In addition to this, 2 ml of reagent X was added (while grinding) to form a homogenous mixture and allowed to thaw completely by grinding intermittently. Then the whole solution was transferred to 2 ml micro-centrifuge tubes and kept without disturbing for 5 min. at optimum room temperature of 26 ± 2°C. To the solution, 0.2 ml of chloroform was added to each of the tubes and allowed to vortexes gently for few seconds and again kept without disturbing for 10 min. at room temperature. The whole mixture was subjected to centrifugation at 5000 rpm for 10 min. under 4°C and the resultant upper aqueous phase was transferred into a fresh micro-centrifuge tube. About 0.6 volumes of isopropanol was added to it and vortexes for few seconds and kept undisturbed for 10 min. After high speed centrifugation at 13,000 rpm for 10 min. under 4°C, the supernatant was discarded, and the obtained pellet deserve to be rich in RNA content was further washed with 70% ethanol, air dried and diluted in 50 µl of DEPC-treated water. The quality of the extracted RNA content was assessed by using 1.5% Agarose gel electrophoresis.

Preparation of reagent X: Phenol saturated in Tris (hydroxymethyl) aminomethane buffer with a pH of 6.7. Along with this, 0.1% (w/v) of SDS (sodium dodecyl sulphate), 0.32 M (w/v) of sodium acetate and 0.01 M final concentration from a stock of 0.5 M ethlyenediaminetetra acetic acid (EDTA), pH 8.0 was added.

Synthesis of complimentary DNA (cDNA)

The complimentary DNA was synthesized from the obtained RNA by the following method. First strand of complimentary DNA was synthesized by using 3 µl of obtained RNA sample as template along with 1 µl of SMART IV oligonucleotide and CDS/3' PCR primer. Followed by vigorous mixing and incubated at 72°C for 2 min. and

cooled down in ice bath for 2 min. After a short spin, 2 µl of 5x strand buffer, 1 µl of 20 mM DTT, 1 µl of 10 mM dNTP mix and 1 µl of power script RT were gently added, mixed and incubated at 42°C for 1 h and kept stored at -80°C.

PCR amplification

Amplification of both the squalene synthase (SQS) of *Botryococcus braunii* and (SSL-3) Botryococcene synthase genes were performed individually by polymerase chain reaction with a total volume of 30 µl constitutes 15 µl of master mix, 1 µl of 1 µM each of forward and reverse primers and 1 µl of synthesized cDNA template and the rest was made up to 30 µl with Milli-Q water. The PCR conditions for amplifying both the genes were similar for amplification of both the genes. The denaturation reaction was carried out at 94°C for 5 min., followed by annealing at 72°C for 1 min. and extension up to 45 seconds. Final extension temperature was also 72°C for 5 min. Therefore, overall 35 PCR cycles were performed for both the genes separately followed by quality check in 1% agarose gel using 1X TAE buffer with the gel pre-stained by 10 mg/ml of ethidium bromide. The short forward and reverse oligoprimers implemented for PCR amplification of squalene synthase and botryococcene synthase were given in Table 1.

Sequencing of squalene synthase and SSL-3 genes

Qiaquick PCR purification kit (QIAGEN, USA) was used for the purification of amplified PCR products. Both the forward and reverse strands synthesized by forward and reverse primers were sequenced with the use of Big Dye version 3.1. kit (Applied Bio-systems) on an ABI-PRISM 3730 DNA sequencer (Applied Bio-syLLstems). The sequences were assembled using bioinformatics tool Bio-Edit (Version 7.0.9.0), ambiguous sequences were corrected with Chromas (Version 2.01). Both the genes were submitted to NCBI GenBank and accession numbers were retrieved.

In silico modeling and characterization

For In silico modeling, the obtained gene sequences of both the SQS and SSL-3 genes were translated into amino acid sequences by ExPASy online translate server, Swiss Institute of Bioinformatics. The ORF (Open reading frame) showing long chain amino acid sequence was selected for each of the SQS and botryococcene enzymes for the protein prediction. The amino acid sequences of both the SQS and botryococcene synthase enzymes were subjected to NCBI protein BLAST and about 50% conserved sequences were retrieved from NCBI by FASTA file and both the pairwise and multiple sequence alignment were carried out by ClustalW and the Phylogenetic tree was reconstructed based on UPGMA method using MEGA 6 software. JPred 4 online protein prediction server was implemented to predict the protein from the selected ORFs [11]. Similarly, the molecular structure homology modeling of both the SQS and botryococcene synthase enzymes were done in silico by (PS)2-v2: Protein Structure Prediction server [12]. The presence or absence of signal peptides in the target sequences of both the SQS and botryococcene synthase enzymes were predicted by using NetNGly 1.0 online server [13]. The signature sequences are the conserved Motifs present in the target amino acid sequences. Such motifs were analyzed for both the SQS and botryococcene synthase enzymes by ScanProsite online server from Swiss Institute of Bioinformatics [14].

The homology modeling for the SQS and botryococcene synthase enzymes were done in SWISS MODEL server from Swiss Institute of Bioinformatics [15-18]. The template search was performed by blast and HHBlits in SWISS-MODEL template library with ProMod3 Version 1.0.2. The target sequences were searched against primary amino acid sequences present in SMLT by BLAST [19]. Then the templates were chosen based on the highest quality of the model from target-template alignment. The QMEAN score was assessed from the global and per-residue model quality for both the enzyme models [20]. Based on the pairwise interface analysis, the homo-oligomeric structure of the target protein was predicted. The PDB model for both the SQS and botryococcene synthase were retrieved from SWISS-MODEL server and viewed by using UCSF Chimera candidate version 1.11.

The superimpose models were developed for both the SQS and botryococcene synthase enzymes based on the best PDB match from SWISS-MODEL server using UCSF Chimera tool. The ligands for both the enzymes were retrieved from RCBS-PDB online server and matched for the model for SQS and botryococcene synthase in UCSF Chimera. Ramachandran plots were generated for both the SQS and botryococcene synthase enzymes and retrieved by Mol Probity Ramachandran Analysis, Mol probity online server from Duke University [21].

The secondary structure of SQS and botryococcene synthase enzymes were predicted by using PSIPRED v3.3 (Position-specific iterated prediction) secondary structure prediction server [22-24]. The protein domain and putative domain boundaries were predicted for both the enzymes based on DomPred server and PSI-BLAST alignment [25]. The protein-domain homology modeling process was carried for both the enzymes by DomSerf v2.0 based on PSI-BLAST hits [26].

The rapid fold recognition and fold domain recognition of both the SQS and botryococcene synthase enzymes were predicted by pGenTHREADER and DomTHREADER tools respectively [24,27,28]. Based on the fold domain recognition results, the CATH (Class, Architecture, Topology and Homologous superfamily) classification and functional family for both the enzymes were further analyzed and interpreted by CATH/Gene3D v4.2 online server [29,30]. Based on the EC diversity, the functionally similar kind of protein was predicted in UniProtKB. The accession numbers and protein sequences were retrieved for pairwise alignment with SQS and botryococcene synthase enzymes using ClustalW in MEGA 6 software [31] and the conserved sequences were highlighted based on the alignment using BioEdit [32].

Results

The respective NCBI GenBank accession numbers for sqs gene and Botryococcene synthase (ssl-3) gene are MG755329 and MG755330. The phylogenetic tree showing that the sqs gene of *B. braunii* BB1 strain grouped under a single clade of squalene synthase (SQS) of *Botryococcus braunii* (Figure 2). Similarly, the phylogenetic tree for ssl-3 (Botryococcene synthase) gene of *B. braunii* BB1 strain had shown that it grouped under a separate clade of ssl-3 (Botryococcene synthase) (Figure 3). According to the JPred protein prediction, the protein BLAST results suggests that the SQS and Botryococcene synthase (SSL-3) of *Botryococcus braunii* BB1 strain has more hits with the putative Farnesyl transferase, chain D (PDB: 3WCB) and Squalene synthase, chain A (3WEH) respectively. Similarly, (PS)2-v2 protein prediction results that no hits were found for the enzyme SQS enzyme (Squalene synthase) and hence, the Botryococcene synthase (SSL-3) of *Botryococcus braunii* BB1 strain was found highly similar with the chain B of Human squalene synthase (Farnesyl transferase) (PDB: 1EZF) with 34.76% of protein sequence identity.

Signal peptides are signals for transport across membrane which were absent in both the SQS and SSL-3 (Botryococcene synthase)

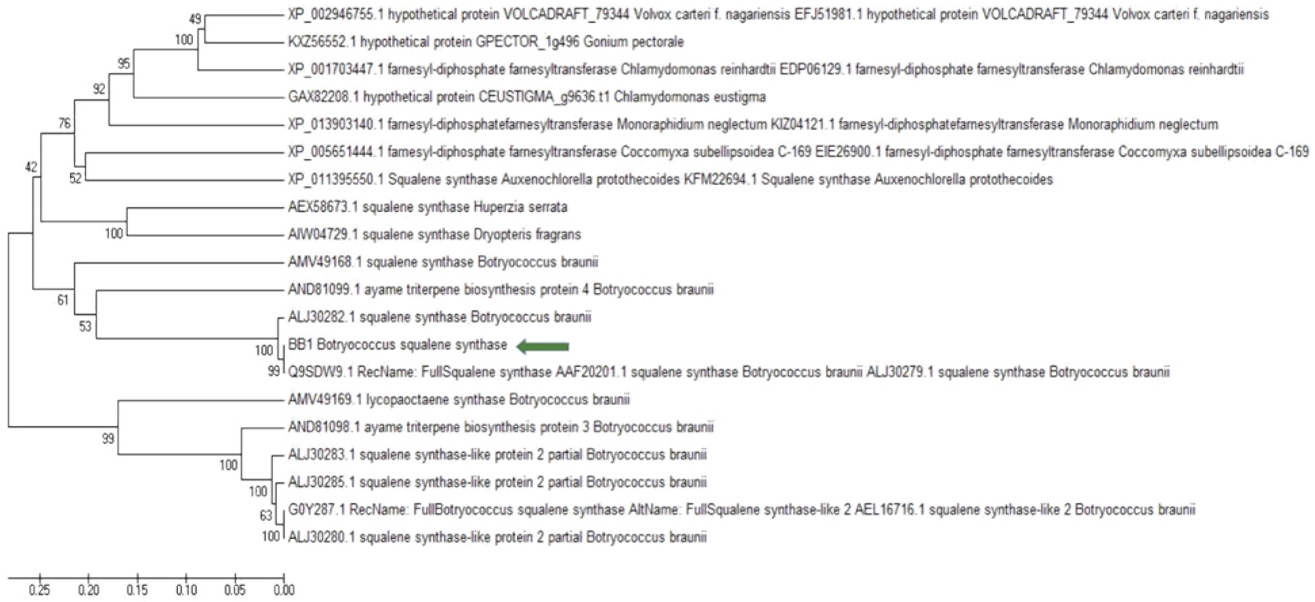


Figure 2: The Phylogenetic tree was constructed from the retrieved protein sequences having high hits with SQS of *B. braunii* BB1 strain based on NCBI BLAST. This Phylogenetic tree was constructed based on the UPGMA statistical method with Bootstrap 500 replications and the gaps were completely deleted. Before Phylogenetic tree construction, multiple sequence alignment was carried out by Clustal W in MEGA6 software tool.

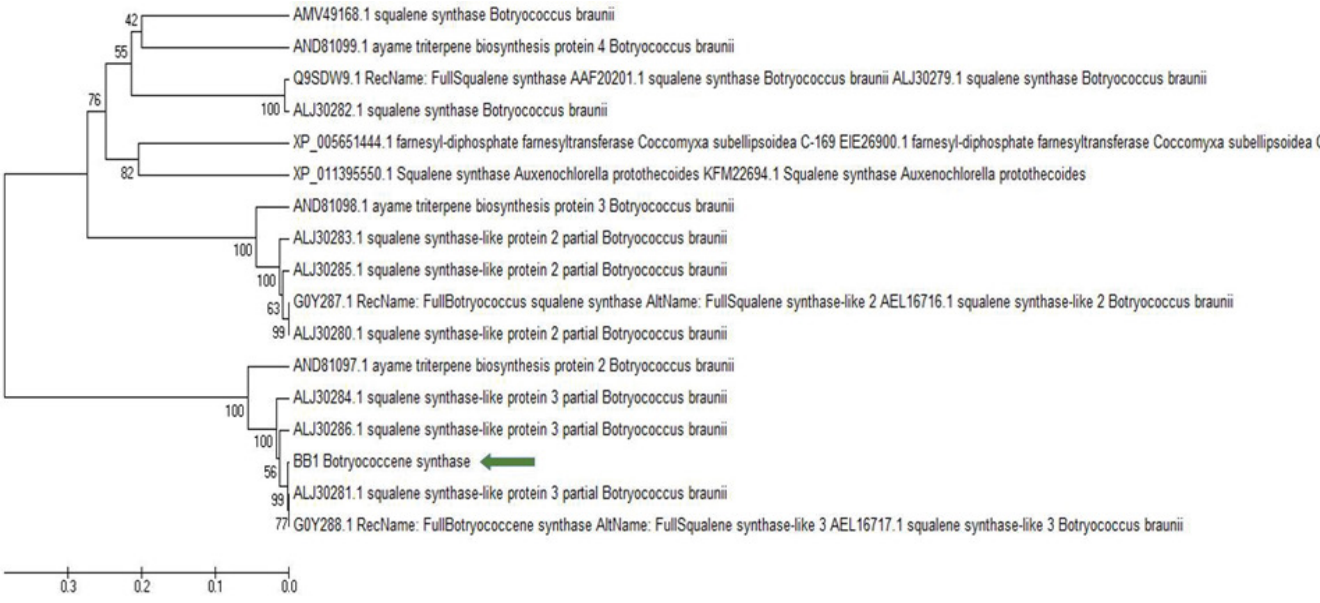


Figure 3: The Phylogenetic tree was constructed from the retrieved protein sequences having high hits with ssl-3 (Botryococcene synthase) of *B. braunii* BB1 strain based on NCBI BLAST. This Phylogenetic tree was constructed based on the UPGMA statistical method with Bootstrap 500 replications and the gaps were completely deleted. Before Phylogenetic tree construction, multiple sequence alignment was carried out by Clustal W in MEGA6 software tool.

enzymes of *B. braunii* BB1 strain as a result from NetNGly. However, N-glycosylation site was found in the SQS enzyme at 310th amino acid residue position. The signature sequence is the most conserved amino acid sequence based on their respective protein function. In our present study, two signature sequences were found in SQS enzyme of *B. braunii* BB1 strain at different positions in the amino acid chain. The first signature sequence was YChyVAGVVGlgLsql found from 172 to 187 amino acid residue position and the second signature sequence was MGlflQkt.NIiRDYfeDinelpapRmFwp between 208 and 236 amino acid residue position. Since, both the signature sequences possess that the signature sequences of SQS enzyme is related to the Squalene and Phytoene synthase enzyme. No such signature sequences were predicted or identified for SSL-3 (Botryococcene synthase) enzyme.

SWISS-MODEL results for SQS enzyme

Based on the SWISSMODEL protein prediction and modeling, the SQS enzyme of *B. braunii* BB1 strain found to have more sequence identity with the chain A of Farnesyl transferase (PDB: 3WCB.1.A)

with about 46.78%. Since, sequence coverage were 0.74 ranges from 36 – 383 amino acid residues. The Farnesyl transferase is a homo-tetramer and matches with the SQS enzyme of *B. braunii* BB1 strain and the ligand was identified based on the most conserved binding site which is 8PH (HYDROGEN[(1R)-2-(3-DECYL-1H-IMIDAZOL-3-IUM-1-YL)-1-HYDROXY-1-PHOSPHONOETHYL] PHOSPHONATE). The conserved binding site constitutes of V49, S50, R51, S52, F53, Y72, R76, V180, G181, L184, M208, G209, L212, N216 these are the amino acid residues which in contact with the ligand 8PH. The QMEAN, Cβ, All atom, Solvation and Torsion values are -3.92, -3.40, -0.71, 1.11 and -3.53 respectively and thus, this predicted model for SQS enzyme of *B. braunii* BB1 strain is the accepted model. The three dimensional structure of SQS enzyme of *B. braunii* BB1 strain was retrieved from SWISSMODEL, showing about 20 helical structures with no sheets in it (Figure 4a). Similarly, the superimposed model showing no major differences in the molecular structures between SQS enzyme and Farnesyl transferase enzyme (PDB: 3WCB) (Figure 4b).

The detailed molecular structure of the ligand 8PH was given in Figure 5a with the molecular formula of C15H30N2O7P2 and molecular weight of 412.36. Figure 5b shows the interactions of amino acids involved in the active site of the enzyme with the ligand 8PH. It is clear that in Farnesyl transferase (PDB: 3WCB), the amino acid residues Y64, F45, S42 and S44 interacts with the phosphate groups of the ligand

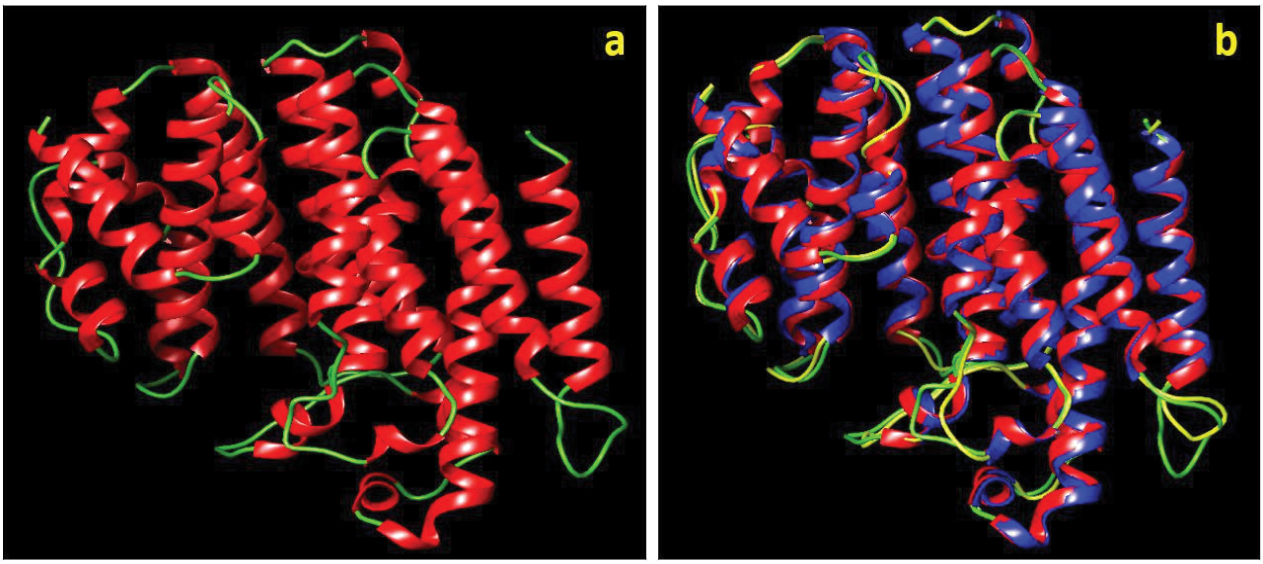


Figure 4: (a) The predicted 3D structure of SQS enzyme of *B. braunii* BB1 strain depicted with helices in red colour and coils in green colour, **(b)** The superimposed 3D structure of both the SQS enzyme of *B. braunii* BB1 strain with the Farnesyl transferase enzyme (PDB: 3WCB) depicted with helices in red and blue colour for SQS and Farnesyl transferase and coils in green and yellow for SQS and Farnesyl transferase respectively.

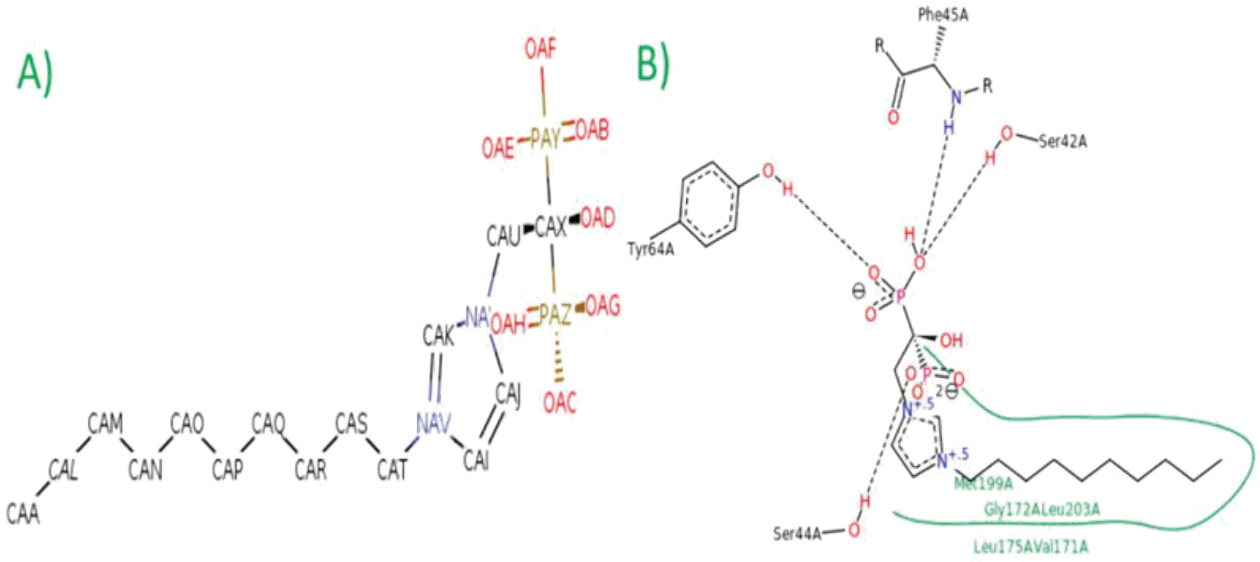


Figure 5: (a) Molecular structure of the Ligand 8pH (HYDROGEN [(1R)-2-(3-DECYL-1H- IMIDAZOL-3-IUM-1-YL)-1-HYDROXY-1-PHOSPHONOETHYL] PHOSPHONATE) and **(b)** the interaction of the ligand with the active site of the enzyme Farnesyl transferase (PDB: 3WCB).

by forming hydrogen bonds whereas, M199, G172, L203, L175 and V171 are the amino acid residues forming hydrophobic interactions with the ligand 8PH. But in the case of SQS enzyme of *B. braunii* BB1 strain, Y72, F53, S50 and S52 interacts with phosphate groups to form hydrogen bonds and M208, G181, L212, L184 and V180 involves in the hydrophobic interactions with the ligand 8 pH (Figure 6).

The Ramachandran plot retrieved for the SQS enzyme of *B. braunii* BB1 strain showing that about 96.5% of the amino acid residues were plotted in the favored regions and 99.1% of amino acid residues were plotted in the allowed regions. There were three outliers found in the Ramachandran map which are 62 ALA, 63 GLN and 61 PRO (Figure 7).

SWISS-MODEL results for SSL-3 (Botryococcene synthase) enzyme

The SWISSMODEL results for SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain reveals that it has high sequence identity with chain A of Farnesyl transferase (PDB: 3WCA.1.A) with 37.50%. The sequence coverage was 0.99 ranges from 8-332 amino acids residues. Even though, the SSL-3 (Botryococcene synthase) enzyme has high hits with Farnesyl transferase based on HHblits, the ligand was not found because the binding site was not conserved. Based on the QMEAN (-3.05), Cβ (-1.91), All atom (-0.58), Solvation (1.52) and Torsion (-3.02) values this model for SSL-3 (Botryococcene synthase) enzyme was accepted a good model. Since, the ligand for this model was found to be Farnesyl thiopyrophosphate (FPS) (S-[(2E,6E)-3,7,11-TRIMETHYLDODECA-2,6,10-TRIENYL] TRIHYDROGENTHIODIPHOSPATE) and Magnesium ion (MG) but the binding sites were not conserved. The 3D structure of SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain retrieved from SWISSMODEL showing about 15 helical structures and the absence of sheet (Figure 8a). The superimpose model for SSL-3 (Botryococcene synthase) enzyme and Farnesyl transferase enzyme (PDB: 3WCA) also showing similar results with the case of the former enzyme (Figure 8b).

The predicted ligand for SSL-3 (Botryococcene synthase) enzyme are Farnesyl thiopyrophosphate (FPS) (Molecular formula: C15H28O6P2S) (Molecular weight: 398.39) (Figure 9a) and a metal ion Magnesium (MG). But, the ligands were not included in the model since the binding sites are not conserved. At the active site of the enzyme Farnesyl transferase (PDB: 3WCA), Y64 and S44 are the amino acid residues forming hydrogen bonds with the two phosphate groups and L175, V171 and L203 are the amino acid residues forming hydrophobic interactions with the ligand (Figure 9b). Similarly, in the case of SSL-3 (Botryococcene synthase) Y68 and S46 are the amino acids may interacts with the two phosphate groups forming hydrogen bonds and L178, V174 and L206 are results in hydrophobic interactions.

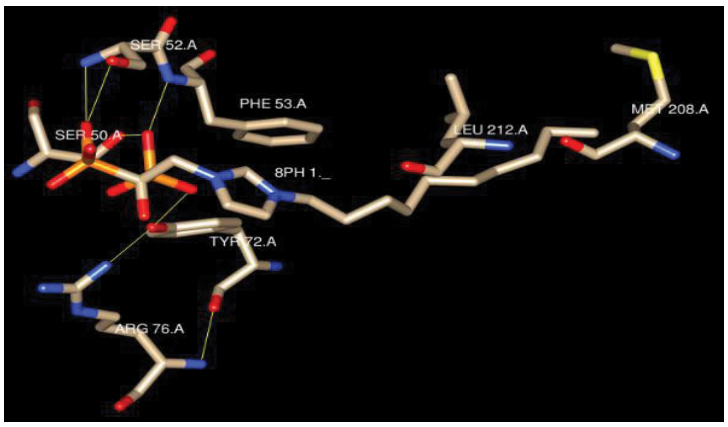


Figure 6: Interactions of amino acid residues with the ligand 8pH at the active site of the enzyme SQS of *B. braunii* BB1 strain.

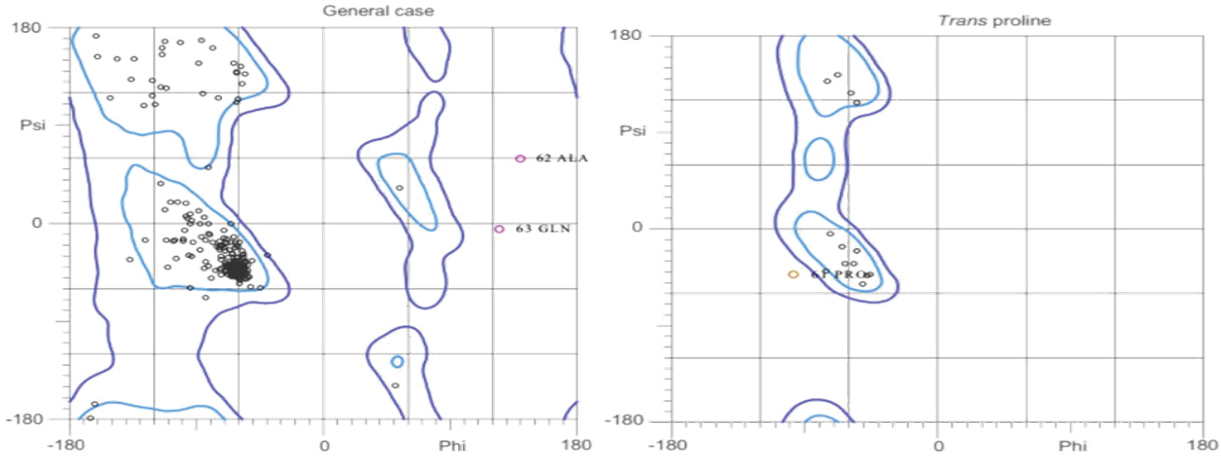


Figure 7: Ramachandran plot analysis of the SQS enzyme of *B. braunii* BB1 strain showing three outliers 63 ALA, 63 GLN and 61 PRO.

The Figure 10 showing the interactions of amino acid residues of the enzyme Farnesyl transferase (PDB: 3WCA) with the ligands Farnesyl thiopyrophosphate (FPS) and a magnesium ion (MG). In which, the interactions of amino acid residues with the ligand FPS is well revealed for SSL-3 (Botryococcene synthase) enzyme. But in the case of magnesium ion (MG), N71, N75 and Q74 are the amino acid residues interact for Farnesyl transferase (PDB: 3WCA). For SSL-3 (Botryococcene synthase) enzyme, D75, D79 and Q78 are in contact with the MG (Mg ion). Therefore, amino acid shift might took place by replacing Asparagine into Aspartic acid at the respective positions for Farnesyl transferase (3WCA) and SSL-3 (Botryococcene synthase) enzyme respectively.

The predicted Ramachandran plot for SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain reveals that 95.0% of all amino acid residues were plotted in the favoured regions. And about 99.1% of the amino acid residues were mapped in the allowed regions. Similar to the case of SQS (Squalene synthase), SSL-3 (Botryococcene synthase) enzyme also has three outliers such as 82 PRO, 83 PRO and 301 ALA (Figure 11).

Based on the PSIPRED secondary protein structure prediction, the SQS (Squalene synthase) enzyme of *B. braunii* BB1 strain consists of about 20 helical structures with a short sheet at 174th and 175th positions and the rest are coils (Figure 12). The PSI-BLAST results reveals that the SQS (Squalene synthase) enzyme of *B. braunii* BB1 strain consists of two domains and the predicted domain boundary location was 331 based on the PSI-BLAST hits of 1872. With this domain boundary prediction, high hits was obtained with the chain B of Human squalene synthase enzyme (PDB: 1EZF) between the query sequence 42 to 319. In the case of SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain, the PSIPRED secondary protein structure predicts that it consists of about 15 helical structures without any sheets and rest are coils (Figure 13). The number of domains present was two and putative domain boundary location is found at 243 based on the PSI-BLAST hits of 2180 for SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain. Based on the domain prediction, high hits was achieved with the chain B of Human squalene synthase enzyme (PDB: 1EZF) for the query sequence between 10 and 332.

The pGenTHREADER is based on highly sensitive fold recognition between protein profile-profile comparison with whole chain library and the net score for SQS enzyme of *B. braunii* BB1 strain was 166.400with chain A of Farnesyl transferase (3WCA) based on the alignment and the alignment score was 795.5 (Figure 14). The DomTHREADER predicts domain based on highly sensitive domain recognition of the query sequence by profile-profile comparison with whole domain library. In this way, the net score was 11.717 for SQS enzyme of *B. braunii* BB1 strain with chain B of Human squalene synthase (PDB: 1EZF) and alignment score was 936.3 (Figure 15). Based on pGenTHREADER results, the SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain has high hits with chain A of Human squalene synthase (PDB:

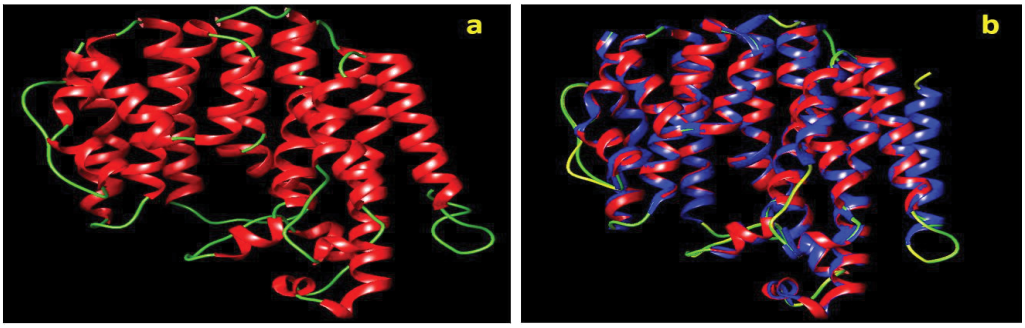


Figure 8: (a) The predicted 3D structure of SSL-3 (Botryococcus synthase) enzyme of *B. braunii* BB1 strain depicted with helices in red colour and coils in green colour, **(b)** The superimposed 3D structure of both the SSL-3 (Botryococcus synthase) enzyme of *B. braunii* BB1 strain with the Farnesyl transferase enzyme (PDB: 3WCA) depicted with helices in red and blue colour for SSL-3 and Farnesyl transferase and coils in green and yellow for SSL-3 and Farnesyl transferase respectively.

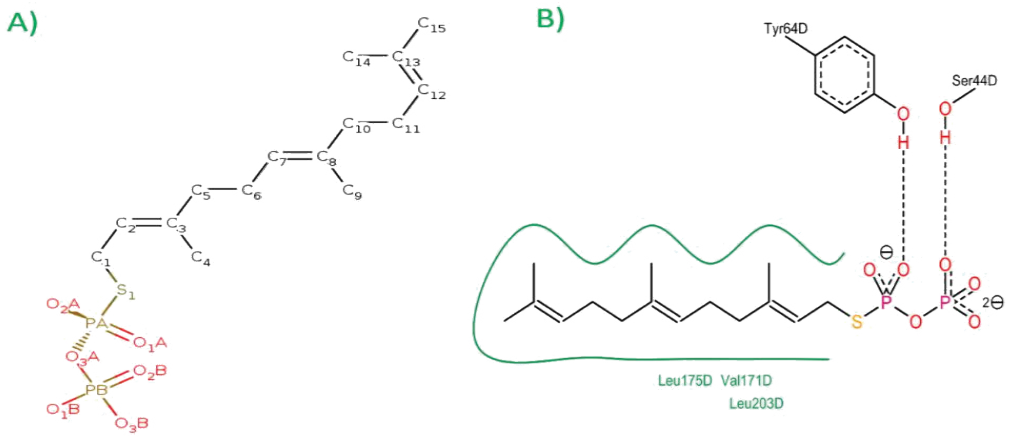


Figure 9: (**a)** Molecular structure of the ligand FPS (Farnesyl thiopyrophosphate) (S- [(2E,6E)-3,7,11-TRIMETHYLDODECA-2,6,10-TRIENYL] Trihydrogen Thiodiphospate) and (**b)** the interaction of the ligand with the active site of the enzyme Farnesyl transferase (PDB: 3WCA).

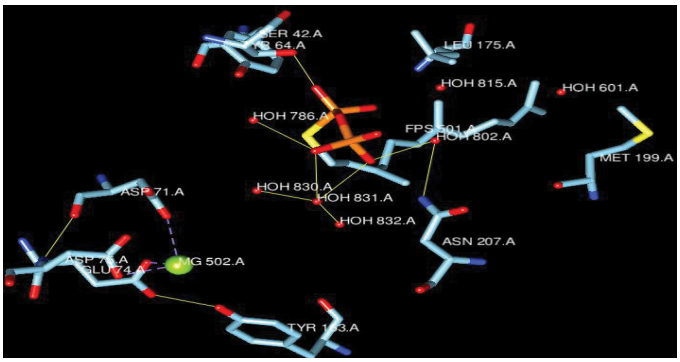


Figure 10: Interactions of amino acid residues with the ligand FPS and MG at the active site of the enzyme SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain.

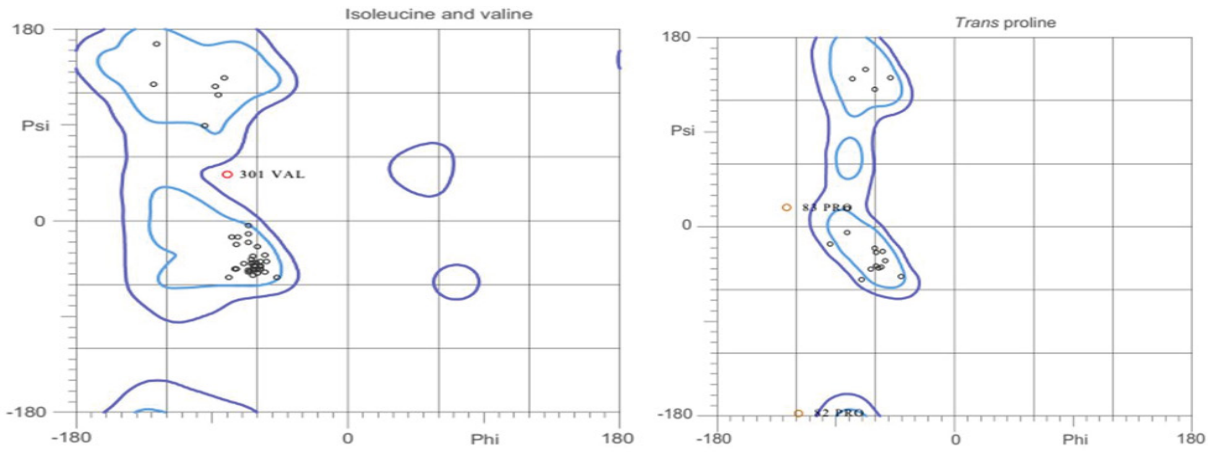


Figure 11: Ramachandran plot analysis of the SSL-3 (Botryococcene synthase) enzyme of B. braunii BB1 strain showing three outliers 82 PRO, 83 PRO and 301 ALA.

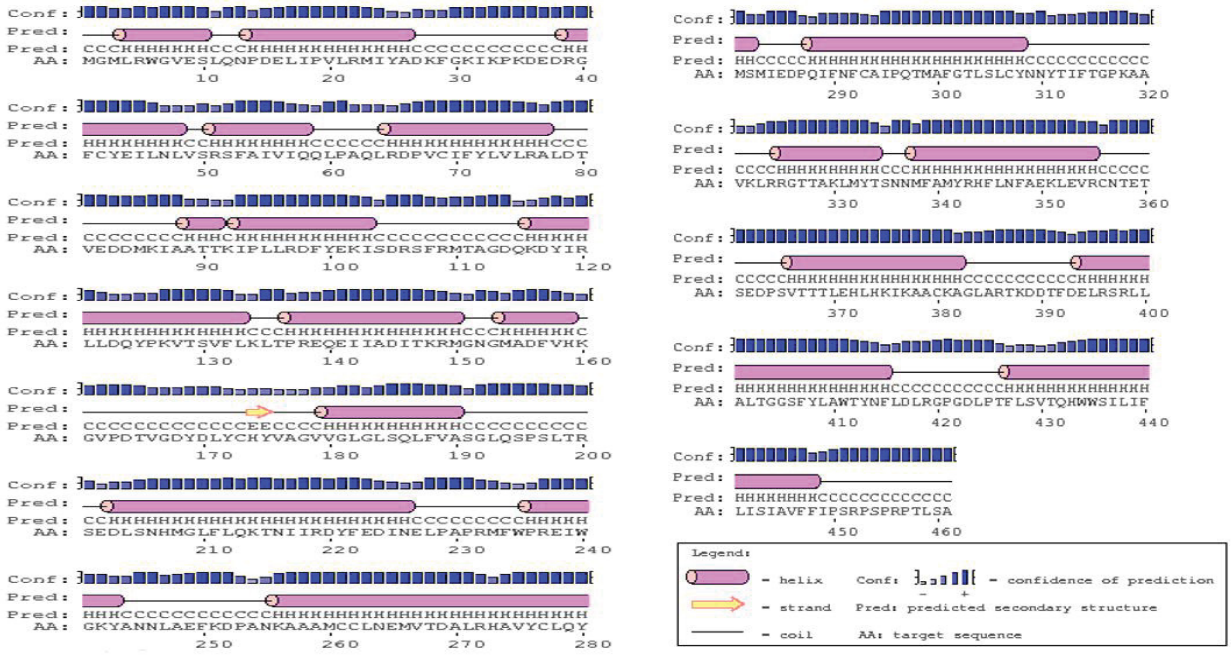


Figure 12: The secondary protein structure of SQS enzyme of *B. braunii* BB1 strain showing 20 helical structures and a very short sheet (174th and 175th positions) and coils.

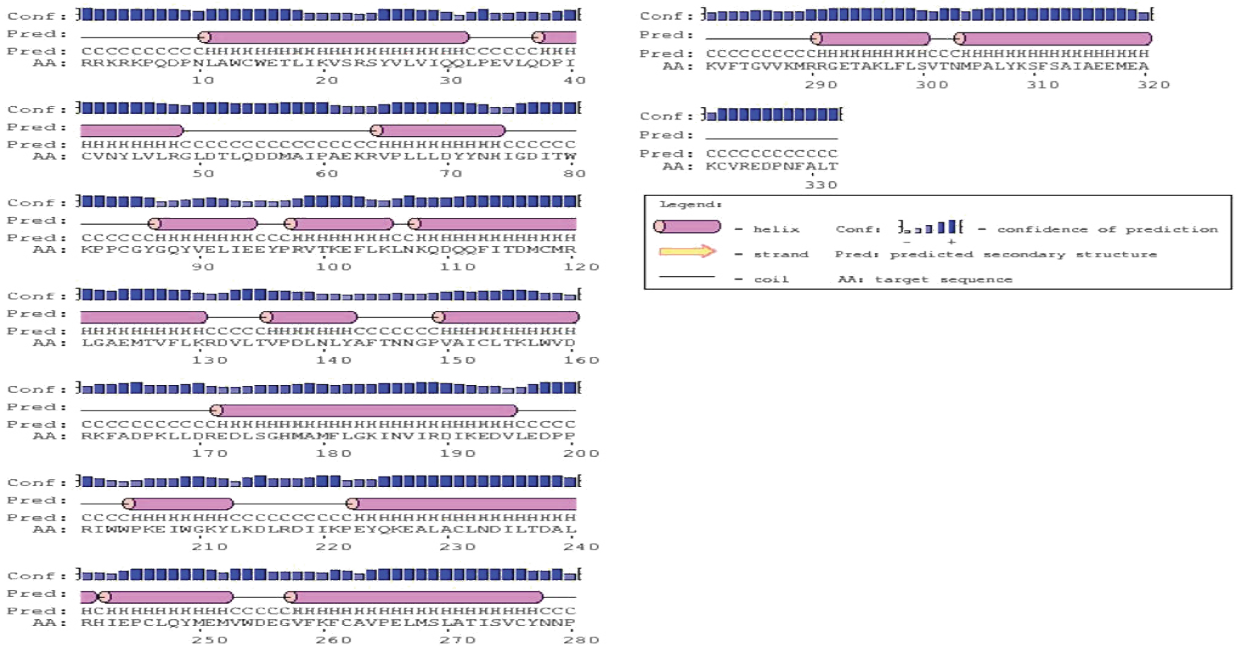


Figure 13: The secondary protein structure of SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain showing 15 helical structures and coils.

pGenTHREADER Scores (alignments displayed against the first of the input sequences)

Conf.	Net Score	p-value	PairE	SolvE	Aln Score	Aln Len	Str Len	Seq Len	View Alignment	SCOP Codes	CATH Codes	Structure	CATH Entry
CERT	166.400	8e-16	-649.2	-23.2	795.5	341	341	461	Display 3wcaA0 Alignment	Search SCOP for 3wcaA	Search CATH for 3wca		CATH Summary
CERT	162.511	2e-15	-634.5	-26.1	765.0	335	335	461	Display 3vj9A0 Alignment	Search SCOP for 3vj9A	Search CATH for 3vj9		CATH Summary
CERT	98.169	6e-09	-257.9	-11.2	467.0	261	263	461	Display 4hd1A0 Alignment	Search SCOP for 4hd1A	Search CATH for 4hd1		CATH Summary
CERT	89.917	4e-08	-298.9	-9.7	400.0	283	284	461	Display 3acxA0 Alignment	Search SCOP for 3acxA	Search CATH for 3acx		CATH Summary
CERT	81.160	3e-07	-313.2	-8.9	344.0	246	247	461	Display 3we9A0 Alignment	Search SCOP for 3we9A	Search CATH for 3we9		CATH Summary
CERT	58.563	6e-05	-445.6	-4.4	139.0	414	901	461	Display 4c0oA0 Alignment	Search SCOP for 4c0oA	Search CATH for 4c0o		CATH Summary
CERT	58.299	7e-05	-376.2	-10.8	142.0	382	523	461	Display 4f52E0 Alignment	Search SCOP for 4f52E	Search CATH for 4f52		CATH Summary

Figure 14: pGenTHREADER scores of SQS enzyme of *B. braunii* BB1 strain.

3VJ9) and the net score was 154.468 and the alignment score was 728 respectively (Figure 16). In the case of DomTHREADER results, based on domain matching the SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain has high hits with the chain B of Human squalene synthase (PDB: 1EZF) with the net score value of 11.07 and alignment score was 313 (Figure 17).

Based on the DomTHREADER results, both the SQS and SSL-3 (Botryococcene synthase) enzymes of *B. braunii* BB1 strain corresponds to a similar kind of domain with respect to Squalene synthase. While analyzing the CATH (Class, Architecture, Topology and Homologous superfamily) classification, both the enzymes are classified under a Superfamily group called Farnesyl diphosphate synthase (CATH Code: 1.10.600.10) and within that both the enzymes belongs to the Functional family Squalene synthase 2 (Functional Family Code: 9540) (Figure 18). Based on CATH classification, both the SQS and SSL-3 (Botryococcene synthase) enzymes of *B. braunii* BB1 strain are belongs to the class Mainly Alpha, Architecture Orthogonal Bundle, Topology Farnesyl diphosphate synthase and Homologous superfamily Farnesyl diphosphate synthase.

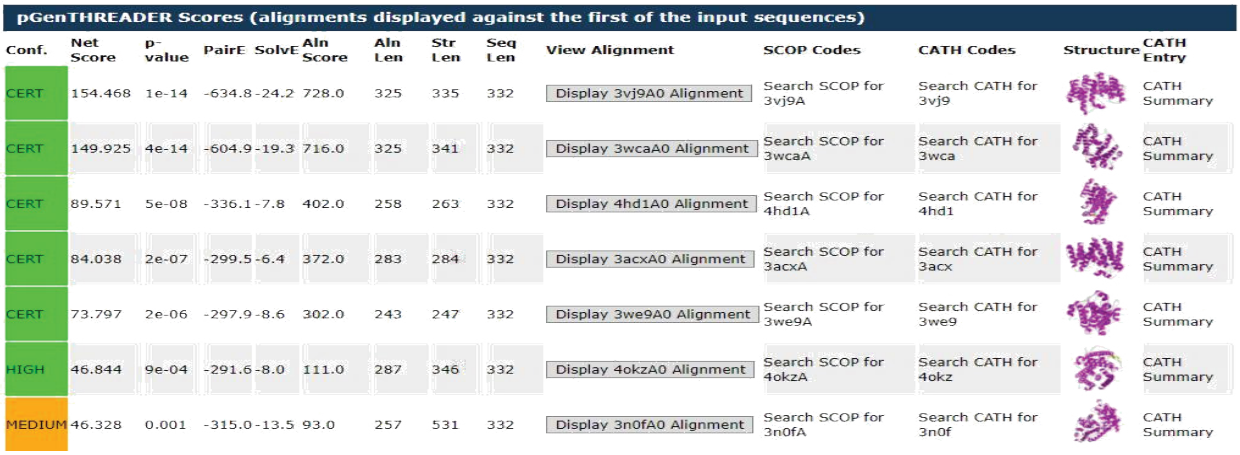
Based on the Functional family of Squalene synthase 2 (Code: 9540) under the CATH of Farnesyl diphosphate synthase (Code: 1.10.600.10), about 17 GO (Gene Ontology) diversity were obtained. Among them squalene synthase (SQS) (GO: 0051996) alone engulfs about 9.3% under molecular function (35.2%) and SSL-3 (Botryococcene synthase) classified under lipid metabolic process (GO: 0019216) (1.9%) of Biological process (29.6%) (Figure 19). Similarly, among six EC (Enzyme Commission) diversity squalene synthase and presqualene diphosphate synthase (SSL-1) distinguished under transferases and the botryococcus

DomTHREADER Scores

Displaying only the top 30 results, full results can be found via the Downloads tab

Conf.	Net Score	p-value	PairE	SolvE	Aln Score	Aln Len	Str Len	Seq Len	Domain Start	Domain End	Domain Code	View Alignment	Structure	CATH Entry
CERT	11.717	2e-08	-552.1	-21.3	936.3	324	324	461	33	380	1ezfB00	Display 1ezfB00 Alignment		CATH Summary
MEDIUM	2.623	6e-04	-409.3	-18.3	95.0	387	1146	461	1	461	1u6gC00	Display 1u6gC00 Alignment		CATH Summary
MEDIUM	2.553	7e-04	-0.0	0.1	96.7	10	73	461	450	461	1qrvA00	Display 1qrvA00 Alignment		CATH Summary
MEDIUM	2.267	1e-03	-0.0	-0.6	91.7	10	164	461	451	461	3h4tA02	Display 3h4tA02 Alignment		CATH Summary
LOW	1.963	0.001	-0.0	-0.1	85.5	13	115	461	448	461	3iiqA01	Display 3iiqA01 Alignment		CATH Summary
LOW	1.940	0.001	-411.4	-9.2	82.0	412	744	461	1	461	2qnaA00	Display 2qnaA00 Alignment		CATH Summary
LOW	1.809	0.002	-0.0	0.0	82.5	9	60	461	451	461	3fghA00	Display 3fghA00 Alignment		CATH Summary
LOW	1.573	0.002	-0.0	-0.6	78.5	12	389	461	450	461	3fwlA02	Display 3fwlA02 Alignment		CATH Summary
LOW	1.468	0.002	-1.2	-0.7	76.6	10	154	461	451	461	2r6fA02	Display 2r6fA02 Alignment		CATH Summary

Figure 15: DomTHREADER scores of SQS enzyme of *B. braunii* BB1 strain.



pGenTHREADER Scores (alignments displayed against the first of the input sequences)

Conf.	Net Score	p-value	PairE	SolvE	Aln Score	Aln Len	Str Len	Seq Len	View Alignment	SCOP Codes	CATH Codes	Structure	CATH Entry
CERT	154.468	1e-14	-634.8	-24.2	728.0	325	335	332	Display 3vj9A0 Alignment	Search SCOP for 3vj9A	Search CATH for 3vj9		CATH Summary
CERT	149.925	4e-14	-604.9	-19.3	716.0	325	341	332	Display 3wcaA0 Alignment	Search SCOP for 3wcaA	Search CATH for 3wca		CATH Summary
CERT	89.571	5e-08	-336.1	-7.8	402.0	258	263	332	Display 4hd1A0 Alignment	Search SCOP for 4hd1A	Search CATH for 4hd1		CATH Summary
CERT	84.038	2e-07	-299.5	-6.4	372.0	283	284	332	Display 3acxA0 Alignment	Search SCOP for 3acxA	Search CATH for 3acx		CATH Summary
CERT	73.797	2e-06	-297.9	-8.6	302.0	243	247	332	Display 3we9A0 Alignment	Search SCOP for 3we9A	Search CATH for 3we9		CATH Summary
HIGH	46.844	9e-04	-291.6	-8.0	111.0	287	346	332	Display 4okzA0 Alignment	Search SCOP for 4okzA	Search CATH for 4okz		CATH Summary
MEDIUM	46.328	0.001	-315.0	-13.5	93.0	257	531	332	Display 3n0fA0 Alignment	Search SCOP for 3n0fA	Search CATH for 3n0f		CATH Summary

Figure 16: pGenTHREADER scores of SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain.

DomTHREADER Scores

Conf.	Net Score	p-value	PairE	SolvE	Aln Score	Aln Len	Str Len	Seq Len	Domain Start	Domain End	Domain Code	View Alignment	Structure	CATH Entry
CERT	11.070	3e-08	-534.3	-20.8	869.0	313	324	332	5	332	1ezfB00	Display 1ezfB00 Alignment		CATH Summary
LOW	1.517	0.002	-325.6	-4.4	74.0	303	1146	332	1	332	1u6gC00	Display 1u6gC00 Alignment		CATH Summary
LOW	1.203	0.003	-0.0	0.2	71.0	11	91	332	1	11	1a0pA01	Display 1a0pA01 Alignment		CATH Summary
LOW	1.158	0.004	-0.0	0.5	70.0	15	41	332	318	332	1bi6H00	Display 1bi6H00 Alignment		CATH Summary
LOW	0.876	0.005	0.1	-0.6	65.5	12	327	332	1	12	3n2tA00	Display 3n2tA00 Alignment		CATH Summary
LOW	0.694	0.006	-46.0	1.3	59.0	88	118	332	163	254	2iubA02	Display 2iubA02 Alignment		CATH Summary
LOW	0.568	0.007	-173.2	-2.7	56.0	302	579	332	1	332	2vglB00	Display 2vglB00 Alignment		CATH Summary

Figure 17: DomTHREADER scores of SSL-3 (Botryococcene synthase) enzyme of *B. braunii* BB1 strain.

squalene synthase (SSL-2) and botryococcene synthase (SSL-3) are classified under oxidoreductases. In which, squalene synthase (EC: 2.5.1.21) alone contributes about 93.9%, presqualene diphosphate synthase (SSL-1) (EC: 2.5.1.103) constitutes 1.8%, botryococcus squalene synthase (SSL-2) (EC: 1.3.1.96) and botryococcene synthase (SSL-3) (EC: 1.3.1.97) engulfs 1.8% and 0.9% respectively (Figure 20). Followed by species diversity, among 813 unique species *Botryococcus braunii* alone constitutes of 1% (Figure 21).

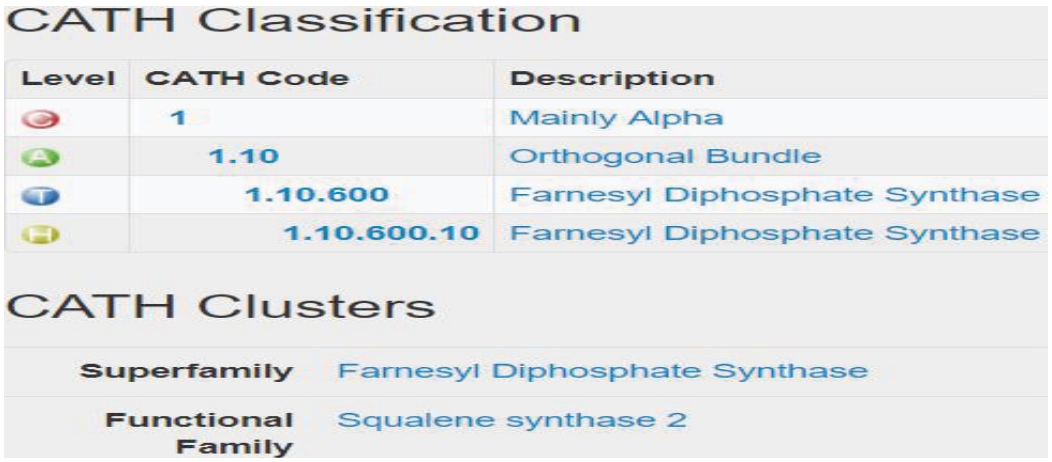


Figure 18: CATH classification of both the SQSand SSL-3 (Botryococcene synthase) enzymes of *B. braunii* BB1 strain.

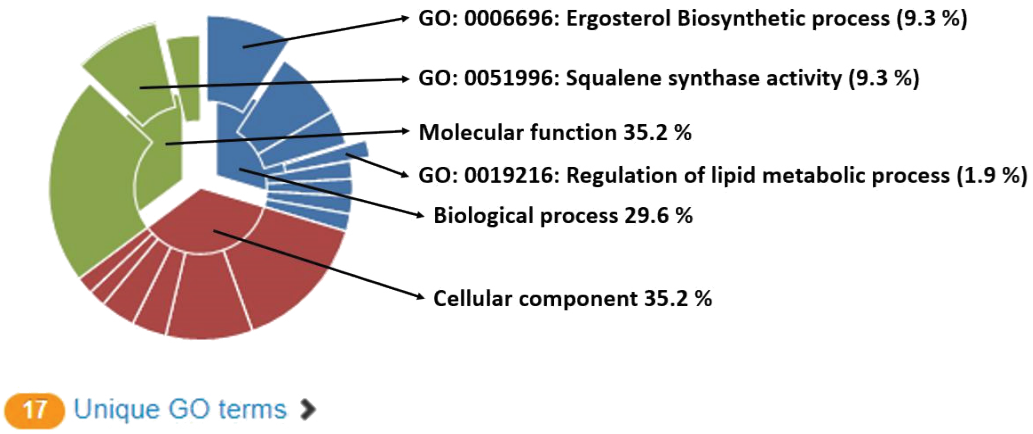


Figure 19: Gene Ontology (GO) diversity of functional family squalene synthase 2 (Functional Family Code: 9540).

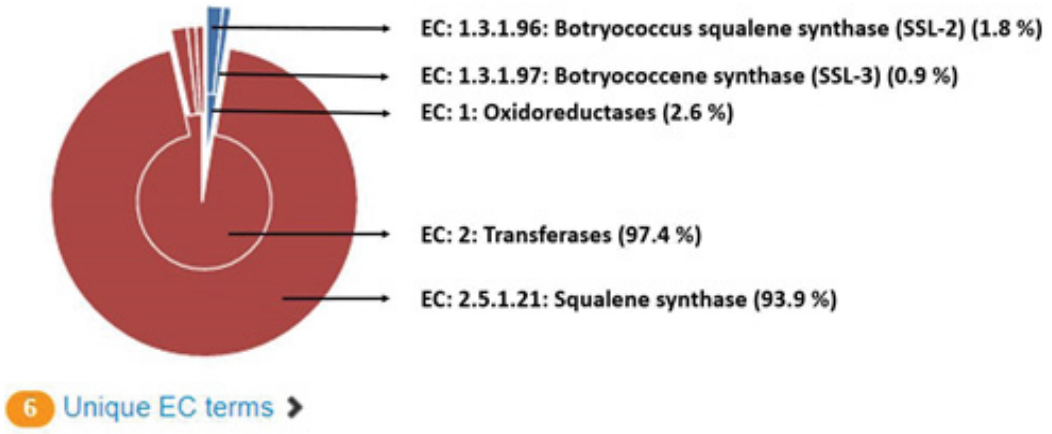


Figure 20: Enzyme Commission (EC) diversity of functional family squalene synthase 2 (Functional Family Code: 9540).

Therefore based on the EC diversity, FASTA file with amino acid sequence of squalene synthase of *Botryococcus braunii* and SSL-3 botryococcene synthase with accession numbers Q9SDW9 and G0Y288 were retrieved. Pairwise sequence alignment were performed in ClustalW with the respective enzymes of *B. braunii* BB1 strain and the results clearly indicates that squalene synthase (SQS) of *B. braunii* of BB1 strain has 100% conserved sites with SQS of *B. braunii* Q9SDW9 (Figure 22). Similarly, botryococcene synthase enzyme also aligned and

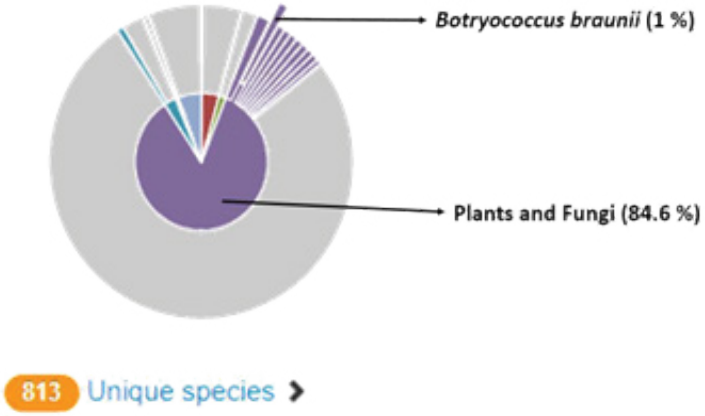


Figure 21: Species diversity of functional family squalene synthase 2 (Functional Family Code: 9540).

```
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            10         20         30         40         50
sp|Q9SDW9|   MGMLRWGVES LQNPDELIPV LRMIYADKFG KIKPKDEDRG FCYEILNLVS
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            60         70         80         90        100
sp|Q9SDW9|   RSFAIVIQQL PAQLRDPVCI FYLVLRALDT VEDDMKIAAT TKIPLLRDFY
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           110        120        130        140        150
sp|Q9SDW9|   EKISDRSFRM TAGDQKDYIR LLDQYPKVTS VFLKLTPREQ EIIADITKRM
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           160        170        180        190        200
sp|Q9SDW9|   GNGMADFVHK GVPDTVGDYD LYCHYVAGVV GLGLSQLFVA SGLQSPSLTR
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           210        220        230        240        250
sp|Q9SDW9|   SEDLSNHMGL FLQKTNIIRD YFEDINELPA PRMFWPREIW GKYANNLAEF
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           260        270        280        290        300
sp|Q9SDW9|   KDPANKAAAM CCLNEMVTDA LRHAVYCLQY MSMIEDPQIF NFCAIPQTMA
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           310        320        330        340        350
sp|Q9SDW9|   FGTLSLCYNN YTIFTGPKAA VKLRRGTTAK LMYTSNNMFA MYRHFLNFAE
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           360        370        380        390        400
sp|Q9SDW9|   KLEVRCNTET SEDPSVTTTL EHLHKIKAAC KAGLARTKDD TFDELRSRLL
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           410        420        430        440        450
sp|Q9SDW9|   ALTGGSFYLA WTYNFLDLRG PGDLPTFLSV TQHWWSILIP LISIAVFFIP
BB1 Botryo  .......... .......... .......... .......... .......... ..........

                    ....|....| .
                           460
sp|Q9SDW9|   SRPSPRPTLS A
BB1 Botryo  .......... .
```

Figure 22: Pairwise sequence alignment of squalene synthase of *Botryococcus braunii* retrieved from UniProt with accession number Q9SDW9 and SQS of *B. braunii* BB1 strain. The pairwise sequence alignment was carried out by ClustalW in MEGA 6 software tool with BLOSUM as Protein weight matrix. The 100% conserved sites were depicted a dots (.) and gaps were denoted as hyphen (-).

```
                    ....|....| ....|....| ....|....| ....|....| ....|....|
                            10         20         30         40         50
G0Y288_BOC      MKLREVLQHP GEIIPLLQMM VMAYRRKRKP QDPNLAWCWE TLIKVSRSYV
BB1_Botryo      R--------- ---------- -----..... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....|
                            60         70         80         90        100
G0Y288_BOC      LVIQQLPEVL QDPICVNYLV LRGLDTLQDD MAIPAEKRVP LLLDYYNHIG
BB1_Botryo      .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....|
                           110        120        130        140        150
G0Y288_BOC      DITWKPPCGY GQYVELIEEY PRVTKEFLKL NKQDQQFITD MCMRLGAEMT
BB1_Botryo      .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....|
                           160        170        180        190        200
G0Y288_BOC      VFLKRDVLTV PDLDLYAFTN NGPVAICLTK LWVDRKFADP KLLDREDLSG
BB1_Botryo      .......... ...N...... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....|
                           210        220        230        240        250
G0Y288_BOC      HMAMFLGKIN VIRDIKEDVL EDPPRIWWPK EIWGKYLKDL RDIIKPEYQK
BB1_Botryo      .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....|
                           260        270        280        290        300
G0Y288_BOC      EALACLNDIL TDALRHIEPC LQYMEMVWDE GVFKFCAVPE LMSLATISVC
BB1_Botryo      .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ....|....| ....|....|
                           310        320        330        340        350
G0Y288_BOC      YNNPKVFTGV VKMRRGETAK LFLSVTNMPA LYKSFSAIAE EMEAKCVRED
BB1_Botryo      .......... .......... .......... .......... ..........

                    ....|....| ....|....| ....|....| ...
                           360        370        380
G0Y288_BOC      PNFALTVKRL QDVQALCKAG LAKSNGKVSA KGA
BB1_Botryo      ......---- ---------- ---------- ---
```

Figure 23: Pairwise sequence alignment of Botryococcene synthase (SSL-3) retrieved from UniProt with accession number G0Y288 and Botryococcene synthase (SSL-3) of *B. braunii* BB1 strain. The pairwise sequence alignment was carried out by ClustalW in MEGA 6 software tool with BLOSUM as Protein weight matrix. The 100% conserved sites were depicted a dots (.) and gaps were denoted as hyphen (-).

reveals that 99.69% of amino acid sequence was similar. Since, there is a single amino acid shift was took place at the 164th amino acid residue position where aspartic acid was changed into asparagine in the botrococcene synthase of *B. braunii* BB1 strain (Figure 23).

Discussion

The Botryococcene synthase is one among the type of squalene synthase found isolated as squalene synthase-like genes in the green microalga *Botryococcus braunii* race B [5]. The SSL-1 and SSL-2 catalyzes the biosynthesis of PSPP and bifarnesyl ether respectively, hence, SSL-3 does not directly exploit FPP as a substrate. But, when under combinations SSL-1 and SSL-3 together synthesize botryococcene and SSL-1 and SSL-2 in combination synthesize squalene [5].

In our present investigation, the sqs gene and ssl-3 gene of *B. braunii* BB1 strain are grouped under a single clade of squalene synthase (sqs) gene and botryococcene synthase ssl-3 genes of the green alga *Botryococcus braunii* respectively. Based on protein homology modeling and prediction, the SQS and Botryococcene synthase (SSL-3) of *Botryococcus braunii* BB1 strain has more hits with Farnesyl transferase, chain D (PDB: 3WCB) and Squalene synthase, chain A (3WEH) and chain B of human squalene synthase (PDB: 1EZF) respectively. Absence of signal peptides is obviously proven in both the enzyme structure. Signature sequences are the most conserved domain found in the enzymes and it was found that two signature sequences for Squalene and Phytoene synthases 1 and 2 were found in SQS enzyme of *B. braunii* BB1 strain, from 172 to 187 amino acid residue positions and the second signature sequence was between 208 and 236 amino acid residue positions. In contrast to this result, no such signature sequences were present in SSL-3, botryococcene synthase enzyme.

The sequence homology model prediction results revealed that SQS has high sequence identity with chain A of Farnesyl transferase (PDB: 3WCB.1.A) and the predicted ligand was 8PH. The active site of enzyme

Gene (Accession no)	Primer sequence (5'-3')
Squalene synthase of *Botryococcus braunii*	F: TTGGCTCATCGCAATCCAAC
	R: TGATAGGGAGGGGTGGTGAA
Squalene synthase- like 3 (SSL-3) (Botryococcene synthase) (HQ585060)	F: TCGGGAAGTCTTGCAGCACC
	R: AAGCACCCTTAGCTGAAACCTT

Table 1: Oligo primer details used for PCR amplification.

with ligand was well discussed in results part in comparison with the former enzyme farnesyl transferase (PDB: 3WCB). But in the case of SSL-3, botryococcene synthase enzyme the sequence identity was low when compared with the SQS enzyme and predicted ligand was FPS and Mg ion. There is some dissimilarity at the active site of the SSL-3 enzyme and the enzyme farnesyl transferase (PDB: 3WCA). According to Ramachandran plot, both the SQS and SSL-3 of *Botryococcus braunii* BB1 strain has three outliers each.

In coherence with this results, both the SQS and SSL-3 of *B. braunii* BB1 strain shares similar domain corresponds to the functional family of Squalene synthase (FunFam: 9540) belongs to Farnesyl diphosphate synthase (CATH Code: 1.10.600.10) of CATH classification of protein domain. The pairwise sequence alignment reveals that SQS of *B. braunii* BB1 strain has about 100% of conserved sites with SQS of *B. braunii* (Q9SDW9). Whereas SSL-3 of *B. braunii* BB1 strain have shown that it has 99.69% similarity with SSL-3 (G0Y288) with a single amino acid shift (asparagine instead of aspartic acid) at 164th amino acid residue position.

The accumulation of oil is seen abundant in both the intracellular and extracellular matrices [33], in which race B is rich in wide range of long-chain and cross-linked biopolymers to form polymethylsqualene diols [34]. The most common and abundantly generated triterpenes are di and tetra-methylated botryococcenes in race B of *B. braunii* [35]. Other structural derivatives of both squalene and botryococcene ranges from C31 to C37 and the accumulation vary based on the strains and race in response to culture conditions [36]. The squalene synthase genes have been overexpressed in medicinal plants Panax ginseng [37] and Eleutherococcus senticosus to induce the production of phytosterols and triterpenes. However, botryococcene is targeted for biofuel production and hence, the results from the coexpression studies of SSL-1 and SSL-3 genes both individually and in combination reveals the potential for engineering to benefit large scale biosynthesis of botryococcene (Niehaus et al., 2011). Overexpression studies on fusing both the genes SSL-1 and SSL-3 found effective in synthesizing 100 mg/L of botryococcene in yeast [5].

Therefore, these squalene synthase-like genes, SSL-1, SSL-2 and SSL-3 are key genes which can be genetically engineered for the biosynthesis of botryococcene which will be a renewable and sustainable replacement for fossil based fuel. Henceforth, there is a huge gap in the implementation of genetic engineering tool due to lack of high resolution molecular structural studies such enzymes. It is evident in our study, that both the SQS and SSL-3 enzymes have very low protein homology with human squalene synthase enzyme. Thus, before implementing genetic engineering tool we need to study more about the structure of the enzymes and its active site by high resolution X-ray diffraction studies.

Conclusion

The overall results from our study conclude a valuable suggestion to fulfill the huge gap before employing genetic engineering tool for sustainable production of biofuel. The high resolution X-ray diffraction studies on such enzymes will give a clear outline about the molecular structure of enzyme and its active site towards its substrate specificity. It will very useful for us to genetically improve the enzyme efficiency for large scale production of biofuel.

Conflict of Interest

The authors have no conflict of interest.

References

1. Zhang Z, Metzger P, Sachs JP (2007) Biomarker evidence for the co-occurrence of three races (A, B and L) of *Botryococcus braunii* in El junco Lake, Galapagos. Org Chem 38: 1459-1478.
2. Moldowan JM, Seifert WK (1980) First discovery of botryococcane in petroleum. J Chem Soc Chem Comm 19: 912–914.
3. Mc Kirdy DM, Cox RE, Volkman JK, Howell VJ (1986) Botryococcane in a new class of Australian non-marine crude oils. Nature 320: 57–59.
4. Dayananda C, Sarada R, Kumar V, Ravishankar GA (2007) Isolation and characterization of hydrocarbon producing green alga *Botryococcus braunii* from Indian freshwater bodies. Electron J Biotechnol 10.
5. Niehaus TD, Okada S, Devarenne TP, Watt DS, Sviripa V (2011) Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii*. PNAS 108: 12260-12265.
6. Lee S, Poulter CD (2008) Cloning, solubilization and characterization of squalene synthase from Thermosynechococcus elongatus BP-1. J Bacteriol 190: 3808-3816.
7. Hillen LW, Pollard G, Wake LV, White N (1982) Hydrocracking of the oils of *Botryococcus braunii* to transport fuels. Biotechnol Bioeng 24: 193–205.
8. 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.
9. 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.
10. 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.
11. Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: A protein secondary structure prediction server. Nucleic Acids Res 43: 389-394.
12. Chen CC, Hwang JK, Yang JM (2009) (PS)2-v2: Template-based protein structure prediction server. BMC Bioinformatics 10: 366.
13. Gupta R, Jung E, Brunak S (2004) Prediction of N-Glycosylation sites in human proteins. In Preparation.
14. Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, et al. (2006) Scan prosite: Detection of PROSITE signature matches and pro rule associated functional and structural residues in proteins. Nucleic Acids Res 34: 362-365.
15. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, et al. (2014) SWISS-MODEL: Modeling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42: 252-258.
16. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T (2017) Modeling protein quaternary structure of homo- and hetero oligomers beyond binary interactions by homology. Sci Rep 7: 10480.
17. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS MODEL workspace: A web-based environment for protein structure homology modeling. Bioinformatics 22: 195-201.
18. Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30: 162–173.
19. Altschul SF (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.
20. Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27: 343-350.

21. Lovell SC, Davis IW, Bryan Arendall W, De Bakker PIW, Word JM, et al. (2003) Structure validation by Cα Geometry: φ, ψ and Cβ deviation. Proteins Struct Funct Bioinf 50: 437-450.

22. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED protein analysis Workbench. Nucleic acids Res 41: 349-357.

23. Mc Guffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. Bioinformatics 16: 404–405.

24. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292: 195-202.

25. Marsden RL, Mc Guffin LJ, Jones DT (2002) Rapid protein domain assignment from amino acid sequence using predicted secondary structure. Protein Sci 11: 2814– 2824.

26. Lewis TE, Sillitoe I, Andreeva A, Blundell TL, Buchan DW, et al. (2013) Genome 3D: A UK collaborative project to annotate genomic sequences with predicted 3D structures based on SCOP and CATH domains. Nucleic Acids Res 42: D499-D507.

27. Lobley A, Sadowski MI, Jones DT (2009) pGenTHREADER and pDomTHREADER: New methods for improved protein fold recognition and superfamily discrimination. Bioinformatics 25: 1761-1767.

28. Mc Guffin LJ, Jones DT (2003) Improvement of the GenTHREADER method for genomic fold recognition. Bioinformatics 19: 874–881.

29. Sillitoe I, Lewis TE, Cuff A, Das S, Ashford P, et al. (2015) CATH: Comprehensive structural and functional annotations for genome sequences. Nucleic Acids Res 43: D376–D381.

30. Lam SD, Dawson N, Das S, Lees JG (2016) Gene3D: Expanding the utility of domain assignments. Nucleic Acids Res 44: D404-D409.

31. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 30: 2725-2729.

32. Hall TA (1999) Bio Edit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95-98.

33. Weiss TL, Chun HJ, Okada S, Vitha S, Holzenburg A, et al. (2010) Raman spectroscopy analysis of botryococcene hydrocarbons from the green microalga *Botryococcus braunii.* J Biol Chem 285: 32458–32466.

34. Metzger P, Rager MN, Largeau C (2007) Polyacetals based on polymethyl squalene diols, precursors of algaenan in *Botryococcus braunii* race B. Org Geochem 38: 566–581.

35. Huang Z, Poulter CD (1989) Tetramethyl squalene, a triterpene from *Botryococcus braunii* var showa. Phytochemistry 28: 1467–1470.

36. 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.

37. Lee MH, Jeong JH, Seo JW, Shin CG, Kim YS (2004) Enhanced triterpene and phytosterol biosynthesis in Panax ginseng overexpressing squalene synthase gene. Plant Cell Physiol 45: 976- 984.