STUDIA UNIVERSITATIS BABEŞ-BOLYAI BIOLOGIA, LX, 1, 2015 (p. 23-37)

Identification, isolation and bioinformatic analysis of squalene sinthase-like cDNA fragments in *Botryococcus terribilis* AICB 870 strain

Tiberiu Szőke-Nagy[1,2]✉, Adriana Hegedűs[2], Andreea Baricz[1,2], Cecilia Chiriac[1,2], Edina Szekeres[1,2], Cristian Coman[2] and Nicolae Dragoş[1,2]

SUMMARY. *Botryococcus terribilis* is a freshwater colonial green microalga very similar to *B. braunii,* due to the hydrocarbon biosynthesis and accumulation of those biosynthetic products in the extracellular matrix. The hydrocarbon biosynthesis pathway was intensively studied, especially in differents strains of *B. braunii.* Recent studies revealed the presence of three squalene synthase-like (SSL) enzymes involved in the last steps of hydrocarbon biosynthesis from *B. braunii*. The aim of the study is to identify homologous SSL enzymes to *B. terribilis* AICB 870, a freshwater isolate from Bihor County (Romania), based on new isolated cDNA fragments and bioinformatics analysis of sequenced fragments. Light and fluorescence microscopy observation revealed that AICB 870 strain presents features similar to a *B. terribilis* species, especially simple or branched mucilaginous processes and a high number of lipids vesicles. PCR primers designed using SSL nucleotide sequences from *B. braunii* were successfully used to amplify homologous SSL cDNA fragment in the AICB 870 strain. Bioinformatic analysis of nucleotides and translated amino acid sequences including G+C content, nucleotide frequencies, amino acids frequencies, computed Mw/pI and transmembrane motif prediction showed a high degree of similarity between the SSL identified as pertaining to *Botryococcus braunii* and those generated in the present work. The results of the present study pointed out for the first time the presence of three squalene synthase-like enzymes in a strain of *B. terribilis* species.

Keywords: Bioinformatics analysis, *Botryococcus terribilis,* cDNA, hydrocarbons, squalene synthase-like.

[1] *Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania.*
[2] *Institute of Biological Research (NIRDBS), Cluj-Napoca, Romania.*
✉ ***Corresponding author: Tiberiu Szoke-Nagy**, Faculty of Biology and Geology, Babeş-Bolyai University, 5-7 Clinicilor Street, Cluj-Napoca, Romania, E-mail: tiberiu.szoke@ubbcluj.ro*

Introduction

Botryococcus genus comprises 16 species, of which 13 are currently accepted, including *Botryococcus braunii* Kutzing (Guiry and Guiry, 2015). *B. braunii*, especially strains belonging to race B, has been the most studied species of the genus in respect to hydrocarbon biosynthesis. Recently, beside *B. braunii*, another species pertaining to *Botryococcus* genus, named *B. terribili*s, was reported to accumulate hydrocarbons, mostly botryococcenes (Hegedűs *et al.*, 2014).

Botryococcus braunii is a freshwater, colonial, green microalga, which can synthesize high amounts of hydrocarbons within the cell that are accumulated in the extracellular colonial matrix. The hydrocarbons can constitute between 27 and 86% of its dry weight (Brown *et al.*, 1969).

The *B. braunii* strains can produce various types of hydrocarbons, which have been traditionally classified into three chemical races (A, B and L), according to the hydrocarbon oils synthesized: i) race A produce odd-numbered (C_{23} to C_{33}) n-alkadiene and triene (Largeau *et al.*, 1980; Metzger *et al.*, 1985), derived from fatty acids through the very long-chain fatty acids elongation pathway (Baba *et al.*, 2012); ii) race B produce unsaturated triterpene (squalenes, botryococcenes and their methylated derivatives), having general formula C_nH_{2n-10}, n=30-37 (Metzger *et al.*, 1987; Okada *et al.*, 1995); iii) race L which produce a single type of hydrocarbon known as lycopadiene ($C_{40}H_{78}$) (Metzger *et al.*, 1990). Moreover, recently, a new chemical race was described based on GC/MS analysis named race S, which produces epoxy-n-alkane and saturated n-alkane (Kawachi *et al.*, 2012).

Strains belonging to race B are intensively studied because their hydrocarbons have been converted by standard hydrocracking reactions to combustible fuels, including gasoline, kerosene and diesel (Hillen *et al.*, 1982).

Biosynthesis of hydrocarbons in race B takes place in the presence of isopentenyl diphosphate (IPP), supplied through the mevalonate-independent pathway (Sato *et al.*, 2003). Subsequently, farnesyl diphosphate (FPP) is produced from IPP and its isomer dimethyl diphosphate (DMPP). Further, the hydrocarbon biosynthesis can be described as a two-step reaction mechanism (Poulter, 1990): firstly, two molecules of the FPP form presqualene diphosphate (PSPP) through a head-to-head condensation (Sasiak and Rilling, 1988); secondly, in the presence of NADPH, farnesyl moieties are rearranged in order to form squalene with a C1`-1 linkage or botryococcenes with a C1`-3 linkage, between the farnesyl moieties (Blagg *et al.*, 2002).

The enzymes associated with hydrocarbon biosynthesis have been mostly unknown, until Niehaus *et al.* (2011) isolated and characterized three unique squalene synthase-like (SSL-1, SSL-2 and SSL-3) genes by screening cDNA libraries under low stringency hybridization and by computational screening of *B. braunii* transcriptomic data. According to Niehaus *et al.* (2011) the yeast expression experiments of single

SSL enzymes revealed that the SSL-1 is involved in the biosynthesis of PSPP from two moieties of FPP, the SSL-2 produce especially bisfarnesyl ether and low amounts of squalene, while SSL-3 did not lead to the accumulation of any major product. Subsequently, the coexpression of SSL-1 and SSL-2 revealed the biosynthesis of squalene and significant amounts of bisfarnesyl ether, but interestingly, when SSL-1 and SSL-3 were coexpressed together, botryococcenes biosynthesis occurred. These new findings are contradictory to the hypothesis that only squalene synthase (SQS) catalyzed both reactions from the last steps of the hydrocarbon biosynthesis pathway (Okada *et al.*, 2004).

Although the SQS enzymes were described among different taxa (a short list is detailed in the Material and Methods section), including some species of algae, e.g. *Chlamydomonas reinhardtii* (Merchant *et al*., 2007), *Auxenochlorella protothecoides* (Gao *et al*., 2014), *Bathycoccus prasinos* (XP_007512409.1), the SSL enzymes were solely described and studied in *Botryococcus braunii* (Niehaus *et al*., 2011; Bell *et al*., 2014). Thus, the aim of this study was to determine the presence of homologous squalene synthase-like enzymes in *B. terribilis* AICB 870, an algal strain isolated from Bihor County (Romania), study based on direct sequencing of cDNA fragments, synthesized from purified algal RNA and bioinformatics analyses. The importance of searching for SSL enzymes in other (*Botryococcus*) species is given by the possibility to compare the hydrocarbons biosynthesis pathways in order to: i) find one strain with a higher growth rate which can produce large amounts of hydrocarbons; ii) compare the SSL gene expression levels and correlate them with the amount of hydrocarbons biosynthesized; and iii) based on the previously mentioned reasons, it is possible to identify a *Botryococcus* strains with a particular set of SSL gene, which can be further overexpressed in bacteria, yeast or other organisms with higher growth rate in order to obtain large amounts thus efficiently increasing the production of hydrocarbons.

Material and methods

Strain culture

The AICB 870 strain was isolated from Cristur fishpond, Bihor County and AICB 874 strain was isolated from Tăureni fishpond, Mureş County, Romania, both algal strains are deposited in the Algal and Cyanobacterial Culture Collection (AICB) at the Institute of Biological Research, Cluj-Napoca, Romania (Dragoş *et al.,* 1997). The algal cultures of *B. terribilis* AICB 870 and AICB 874 were grown on BBM medium, under continuous irradiation of approximately 150 µmol photons $m^{-2}s^{-1}$, a temperature of 25°C±2°C, in 500 ml conical bubbler using a continuous airlift system. The biomass was harvested in the exponential growth phase and aliquoted for subsequent molecular analysis.

Light and fluorescence microscopy

Light and fluorescence microscopy was performed using a Nikon TE-2000 Eclipse microscope equipped with a Nikon D90 photo camera. For hydrocarbon visualization, cells were collected by centrifugation and stained using 1 µl of Nile red (1 mg/ml stock solution dissolved in acetone). Samples were kept in the dark for 5 min, diluted with 1 ml BBM medium and centrifuged 1 min at 10,000 rpm. The rinsing process was repeated three times, in order to remove the excess of dye, followed by fluorescence microscopy analysis (Weiss *et al*., 2012).

RNA Isolation and cDNA synthesis

In order to obtain sufficient quantities of high quality RNA, seven commercially available kits (*Innu PREP RNA Kit* - Analytik Jena, Germany; *Direct-zol™ RNA MiniPrep* - Zymo Research, USA; *ZR RNA MicroPrep™* - ZymoResearch, USA; *ZR Plant RNA MiniPrep™* - Zymo Research, USA; *SV Total RNA Isolation System* - Promega, USA; *TRIzol®* - Ambion-Life Technologies, USA; *Isolate Plant RNA/RNA Kit* - Bioline, UK) and two protocols (Kim *et al.,* 2012; Ghawana *et al.,* 2011), were tested on the AICB 874 strain (chemical race A), and then the best protocol was used on the AICB 870 strain. This RNA extraction strategy was approached due to the following reasons: i) the AICB 874 strain has a simplified cellular organization because colonies rarely occur and cell walls are less rigid; ii) the doubling time of AICB 874 is much higher, thereby cell yields is higher than at the AICB 870; iii) the AICB 874 strain do not have the extracellular matrix, thereby RNA lysis/extraction buffer can disrupt the cells much better than the other strain. All tested kits and protocols were performed according to manufacturer`s and authors instructions. Subsequently, the most efficient protocol was used to extract total RNA from the AICB 870 strain.

RNA electrophoresis was performed using 1.2% formaldehyde denaturing agarose gel in order to verify the integrity of RNA (Farrell, 2010). RNA quantification was performed on a NanoDrop 2000 spectrophotometer (ThermoScientific, USA).

Before cDNA synthesis, RNA samples were treated with DNase. 1U of Turbo DNase Free, (Ambion Europe, UK) was used in order to remove the residual genomic DNA. The cDNA was obtained with a First Strand cDNA Synthesis Kit (ThermoScientific, SUA), using oligo $d(T)_{18}$ primers.

PCR Amplification, Cloning and Sequencing

In order to amplify the SSL cDNA fragments, three pairs of primers were designed for each SSL using Primer-BLAST (Ye *et al.,* 2012) and tested *in silico* with FastPCR 6.4 (Kalendar *et al.,* 2011).

The PCR was carried out with the newly designed primers as follows: each 50 µl reaction volumes containing 1U of Thermo Scientific™ DreamTaq™ DNA Polymerase in 5 µl of the manufacturer`s buffer, 0.25 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM of each primer, and approximately 300 ng of cDNA template. The Touchdown PCR (TD-PCR)

was performed in a TProfessional TRIO Termocycler (Biometra, Germany). TD-PCR cycling condition were: initial denaturation at 94 °C for 2 min followed by 19 cycles of 94 °C for 50 s, T_a+10 °C for 55 s (with an increment of -0.5 °C / cycle), 72 °C for 60 s, followed by another 19 cycles of 94 °C for 50 s, T_a for 55 s, 72 °C for 60 s and a final extension at 72 °C for 2 min.

The PCR products were verified by electrophoresis on a 1% agarose gel in 1 X TAE running buffer, stained with ethidium bromide (1 µg/ml), and visualized on an UVP transilluminator. Subsequently, the PCR products were purified with GeneJet™ Gel Extraction Kit (Fermentas, Canada) and cloned into the pGEM®-T vector (Promega, USA) following the manufacturer's instructions. Plasmids were isolated with GeneJet™ Gel Plasmid MiniPrep Kit (Fermentas, Canada), and sequenced by a commercial company (Macrogen, The Netherlands), with both forward and reverse M13 primers.

Bioinformatic analysis of nucleotides and translated amino acid sequences

The SSL nucleotide sequences obtained in this study were manually corrected for mismatches and ambiguous nucleotides using Chromas Lite 2.1.1 software (Technelysium, Australia). Nucleotide distribution and frequency between our sequences and those from *B. braunii* (HQ585060.1-3) were analyzed with CLC MainWorkbench 7.6 (CLCbio, Danemark).

Using the SSL nucleotide sequences obtained, amino acid sequences were generated using Translate, from ExPaSy (Gasteiger *et al.,* 2005). In order to identify the similarities or differences between SSL sequences from this study and those deposited in the GenBank database (National Center for Biotechnology Information), multiple amino acid sequences alignments were performed using MEGA 6.06 (Tamura *et al.,* 2013). The following homologous and biochemically characterized SQS sequences from GenBank were used for the multiple sequences alignment: *Homo sapiens* - AAA36645.1 (McKenzie *et al*., 1992), *Rattus norvegicus* - AAA42179.1 (Shechter *et al*., 1992), *Mus musculus* - NP_034321.2 (Schechter *et al*., 1994), *Arabidopsis thaliana* - P53799.1 (Nakashima *et al*., 1995), *Solanum chacoense* - AEX26932.1 (Ginzberg *et al*., 2012), *Nicotiana tabacum* - AAB08578.1 (Devarenne *et al*., 1998), *Saccharamyces cerevisiae* - AAA34597.1 (Jennings *et al*., 1991), *Auxenochlorella protothecoides* - KFM22694.1 (Gao *et al*., 2014), *Bathycoccus prasinos* (XP_007512409.1), *Chlamydomonas reinhardtii* - EDP06129.1 (Merchant *et al*., 2007), *Botryococcus braunii* – SSL-1 - G0Y286.1 (Niehaus *et al*., 2011), *Botryococcus braunii* – SSL-2 - G0Y287.1 (Niehaus *et al*., 2011), and *Botryococcus braunii* – SSL-3 - G0Y288.1 (Niehaus *et al*., 2011).

Isoelectric point and molecular weight were predicted using Compute pI/Mw from ExPaSy (Gasteiger *et al.,* 2005). Frequencies of hydrophobic and hydrophilic residues as well as charged residues, were performed with CLC MainWorkbench 7.6 (CLCbio, Danemark). The protein secondary transmembrane domains were identified using TMHMM (Sonnhammer *et al.,* 1998), using default settings.

Results and Discussions

Light microscopy observations of the studied strain showed that *Botryococcus* sp. AICB 870 presents the features of a *Botryococcus terribilis* species, as was described by Komárek and Marvan (1992), FanésTreviño *et al.* (2010), de Queiroz Mendes *et al.* (2012) and Hegedűs *et al.* (2015).

Colonies of AICB 870 strain are ellipsoid or spheroid in shape and sometimes subcolonies can be clearly seen, connected by mucilaginous strands (Fig.1 a-c, black arrow). Simple or branched mucilaginous processes were observed at the periphery of the colony (Fig.1 a-c, white arrow). Occasionally, small lipid droplets can be seen at the surface of the mucilaginous processes. Pyriform-shape cells, more or less radially oriented, are usually completely embedded in a hydrocarbon-rich colonial matrix or slightly emerged from it (Fig. 2 c).

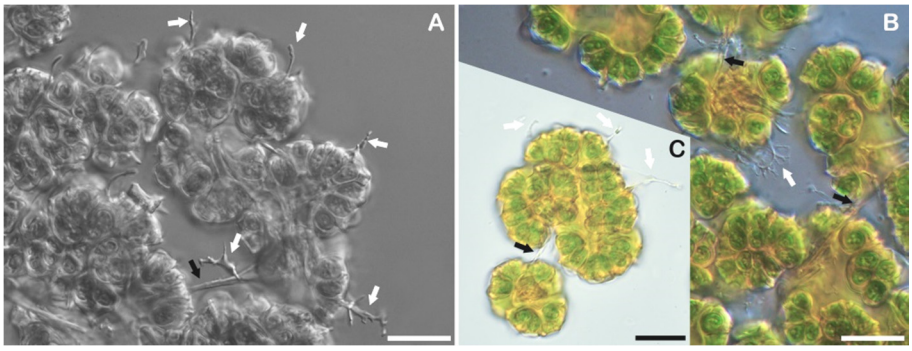

Figure 1. Light micrograph of *Botryococcus terribilis* AICB 870. (A-C) Subcolonies connected by mucilaginous strands (black arrow) and different types of mucilaginous processes (white arrow). Bars = 20 µm.

Colonies stained with a Nile red dye (Fig.2 a-d) showed numerous lipid bodies within the cells (Fig. 2 c-d). Lipid bodies play a key role in hydrocarbon secretion, finally resulting in the accumulation of hydrocarbons in the colonial matrix.

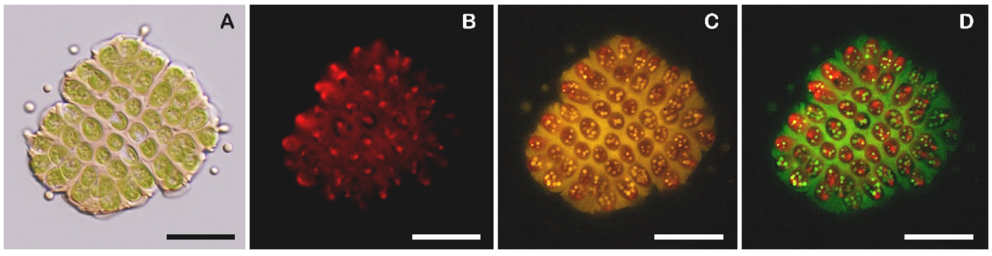

Figure 2. Colony of *Botryococcus terribilis* AICB 870 stained with Nile red dye. (A) Color DIC microscopy image. (B) Chlorophyll autofluorescence - channel. (C) Nile red stained colony. (D) Merged chlorophyll autofluorescence and Nile red stained colony (false colored green). Bar = 20 µm.

RNA Isolation

The results of RNA extraction are presented in Table 1. The RNA yield obtained with the available kits (sample A-G) varied between 9.4 and 175.9 ng/μL, with a very high $A_{260/280}$ ratio but with a low $A_{260/230}$ ratio. Quantitatively, the highest RNA concentration was obtained in the samples H (867.2 ng/μL) and I (3228.4 ng/μL) from AICB 874, further protocol H and I were tested at the AICB 870 (samples H* and I*).

RNA electrophoresis yielded faint bands in samples A-G (except sample D and F) and highly strong bands in samples H, I, and I*. RNA isolated from sample H* was partially or totally degraded. DNA contamination was confirmed in all samples.

The low quality and quantity of the extracted RNA performed with the commercially available kits may be due to the lacking of lysis buffer in order to disrupt the cells and extract the total RNA. This disadvantage was removed from the tested protocols because phenol was used as lysis and extraction buffer combined with silica beads and freeze/thaw.

Table 1.

The RNA yield (ng/μL), $A_{260/280}$ and $A_{260/230}$ ratios of all tested protocols.

Kit / Protocol	***Sample***	***ng/μL***	***$A_{260/280}$***	***$A_{260/230}$***
Innu PREP RNA Kit	A	16.0	2.18	0.92
Direct-zol™ RNA MiniPrep	B	57.9	1.94	1.57
ZR RNA MicroPrep™	C	129.3	1.88	0.78
ZR Plant RNA MiniPrep™	D	175.9	2.02	1.54
SV Total RNA Isolation System	E	9.4	1.65	0.35
TRIzol®	F	162.4	2.00	1.29
IsolatePlant RNA/RNA Kit	G	18.7	1.99	0.82
Kim *et al.*, (2012)	H	867.2	2.04	2.45
Ghawana *et al.*, (2011)	I	3228.4	1.98	1.92
Kim *et al.*, (2012)	H*	229.6	2.02	2.28
Ghawana *et al.*, (2011)	I*	966.1	1.99	1.51

The great advantage of the last tested protocol (Ghawana *et al.*, 2011) is the ability to extract very large amounts of RNA, but there are several disadvantages: residual DNA present in RNA probes and toxicity of the phenol used as extraction agent.

PCR Amplification of cDNA fragments

Starting from the SSL sequences published by Niehaus *et al.*, (2011), for each SSL three pairs of primers were designed in this study using Primer-BLAST (Ye *et al.,* 2012). The primers are detailed in Table 2, with regards to their 5`-3` sequence, T_m and expected length of amplicons. PCR products expected lengths were predicted *in silico* using FastPCR 6.4 (Kalendar *et al.,* 2011).

Using the newly designed primers, four SSL fragments were successfully amplified (Fig.3) using the cDNA as template. For SSL1, a 1200 bp fragment was amplified with the primer pair no. 2 and a 1000 bp fragment with the primer pair no. 3. For SSL2 only one fragment was amplified using the primer pair no. 6 with an approx. length of 1300 bp. A single fragment of approx. 1000 bp in length was amplified for SSL3 using the primer pair no. 9.

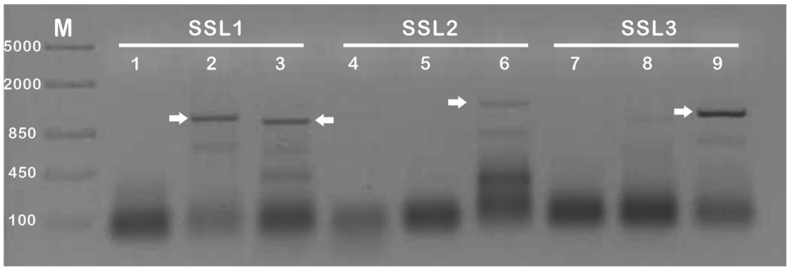

Figure 3. 1% agarose gel electrophoresis for the PCR products obtained through touchdown PCR (white arrow). Marker: FastRuler Middle Range.

Table 2.

The primers pairs used in this study to amplify squalene synthase-like 1-3 cDNA fragments from *B. terribilis* AICB 870.

Primer Pair	Primer name	5`-3` Sequence	T_m (°C)	Expected length (bp)
1	SSL1-F	ATGACTATGCACCAAGACCACG	54.8	1212
	SSL1-R	TCACTTGGTGGGAGTTGGGG	55.9	
2	SSL1-F2	ATGACTATGCACCAAGACCACGG	57.1	1169
	SSL1-R3	GAGGGTGTGTCATACTTGGC	53.8	
3	SSL1-70F	CTCCAACTAGCTGCAACAACGT	54.8	1100
	SSL1-R3	GAGGGTGTGTCATACTTGGC	53.8	
4	SSL2-F	ATGGTGAAACTCGTCGAGGTTT	53.0	1398
	SSL2-R	CTACTGCTTGGAAGAAGCAGAG	54.8	
5	SSL2-F2	CAGATGTTGCATAAGACCTACC	53.0	1334
	SSL2-R2	GAAGAAGCAGAGGTGAGCAAGG	56.7	
6	SSL2-50F	ATGTTGCATAAGACCTACCGCG	54.8	1326
	SSL2-R3	AGCAGAGGTGAGCAAGGGAAGG	56.7	
7	SSL3-F	ATGAAACTTCGGGAAGTCTTGC	53.0	1152
	SSL3-R	CTAAGCACCCTTAGCTGAAACC	54.8	
8	SSL3-F2	ATGAAACTTCGGGAAGTCTTGCAGC	57.7	1149
	SSL3-R2	AGCACCCTTAGCTGAAACCTTTCC	57.4	
9	SSL3-40F	CCCTCTCCTGCAAATGATGGTC	56.7	1058
	SSL3-1099R	ATAACGCCTGGACATCCTGAAG	54.8	

The PCR reaction was repeated using the PCR products obtained in the first reaction as templates, purified from agarose gel, cloned in pGEM-T vector and sequenced.

Bioinformatic analysis of nucleotides and translated amino acid sequences

The bioinformatic analysis was difficult due to a very low number of homologous sequences deposited in the GenBank database and related to *B. braunii* or *B. terribilis*.

The nucleotide sequences generated in the present work were identified based on BLAST report. Our sequences showed a percentage of identity with the SSL-1-3 from *B. braunii* (HQ585060.1-3) that varies between 96% (SSL-2) and 98% for SSL-1 and SSL-3, respectively.

The nucleotide frequencies showed few differences between the SSLs in the two different strain analyzed. The G+C content of the analyzed SSL fragments was: 49.7% in *B. braunii,* 49.5% in AICB 870, for SSL-1; 54.0% in *B. braunii,* 54.2% in AICB 870, for SSL-2 and 50.4% in *B. braunii,* 50.9% in AICB 870, for SSL-3. Overall, from this point of view the CDS fragments from the AICB 870 strain, show high similarity to sequences from *B. braunii*.

The nucleotide sequences were translated into amino acid sequences with the Translate tool from ExPaSy (Gasteiger *et al*., 2005) using the standard genetic code.

The amino acids sequences showed high similarity to SSL from *B. braunii*, as it follows: 97% (399 aa out of 410 aa) with SSL-2 (G0Y287.1) and 99% (398 aa out of 403 aa and 346 aa out of 348 aa) with SSL-1 (G0Y286.1) and SSL-3 (G0Y288.1).

Based on amino acids frequencies (Table 3) it can be observed that the frequencies of hydrophobic amino acids residues are at least two times higher than the hydrophilic ones, varying between 0.474 in SSL-1 and 0.52 in SSL-2 and 3.

The calculated Mw, detailed in Table 3, showed slightly few differences for all six analyzed SSLs. The pI showed differences only in the case of SSL-1 (7.12 at SSL-1 from *B. braunii* and 7.96 at SSL-1 from AICB 870). Using sequence alignments, two possible amino acids substitutions were identified which can result in changing of isoelectric point: i) Glu256 from *B. braunii* SSL-1 is replaced by Lys in SSL-1 from *B. terribilis* AICB 870; and ii) Lys263 is replaced by Arg.

Table 3.

The amino acids frequencies, molecular weight (Mw) and isoelectric point (pI) of all SSL fragments from *B. braunii* Showa (G0Y286-8) and AICB 870 (*complete aa sequence, **partial aa sequence, H^+- hydrophilic, H^- - hydrophobic).

Strain	Fragment	Length (aa)	AA frequencies			Mw (kDa)	pI
			H^-	H^+	Others		
B. braunii Showa	SSL-1	403*	0.474	0.268	0.258	45.95	7.12
AICB 870	SSL-1	403*	0.474	0.268	0.258	45.96	7.96
B. braunii Showa	SSL-2	410**	0.527	0.222	0.251	45.81	6.11
AICB 870	SSL-2	410**	0.529	0.224	0.246	45.84	6.11
B. braunii Showa	SSL-3	348**	0.520	0.201	0.279	40.49	5.95
AICB 870	SSL-3	348**	0.520	0.201	0.279	40.49	5.95

Multiple sequence alignments between our sequences and the sequences retrieved from GenBank, revealed four conserved domains, two aspartate-rich motifs and one NADPH binding domain (Fig. 4), a similar situation being previously described by Lee and Poulter (2008) in the case of the squalene synthase from *Thermosynechococcus elongatus* strain BP-1.The most conserved motif was observed in domain I. The motif (CVF[YL]V[LR]AL[DT]VE[DD]) consisted of 16 amino acids residues of which eight residues are perfectly conserved (brackets). Domain I presented the first aspartate-rich motif [DT]VE[DD]. Domain II had the consensus sequences [D]L[Y]CHY VA[G]LVGIG and presented a partial degree of conservation. The second aspartate-rich motif [D]YL[ED] was observed in domain III. Domains I and III are possibly involved in the binding of substrate (FPP in the case of the first 14 sequences and PSPP in the last four sequences) via Mg^{2+} bridging (Lee and Poulter, 2008; Pandit *et al.*, 2000; Gu *et al.*, 1998). Domain IV presents the highest degree of variability from all four domains and could be involved in rearrangement of PSPP to squalenes or botryococcenes. Downstream of domain IV, all sequences present the domain VKIRK which has role in binding of NADPH (Lee and Poulter, 2008).

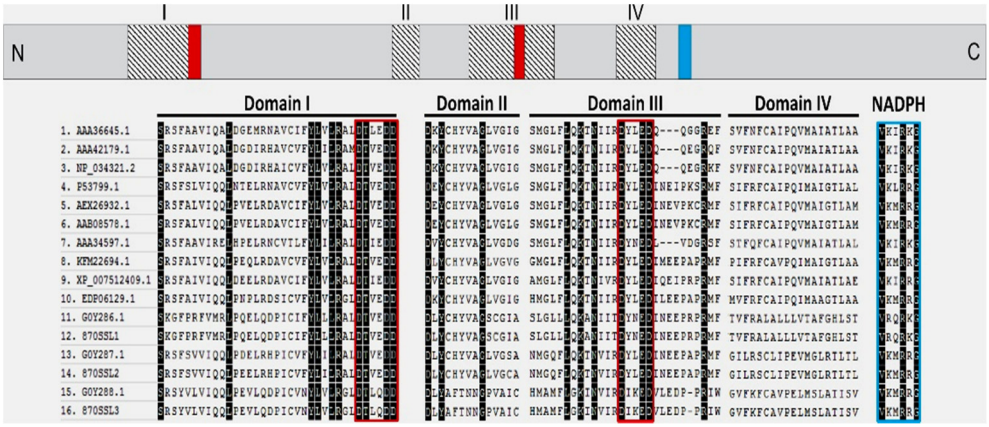

Figure 4. Amino acid sequence alignments of SQS. The conserved domain I-IV, aspartate-rich motifs (red rectangle) and putative NADPH binding domain (blue rectangle) are also depicted.

The 16-aa sequences were checked (mentioned in the Material and Methods section) in order to predict transmembrane motifs (TMM) using TMHMM program which are depicted in Fig. 5 (the identical TMHMM plots are not shown).

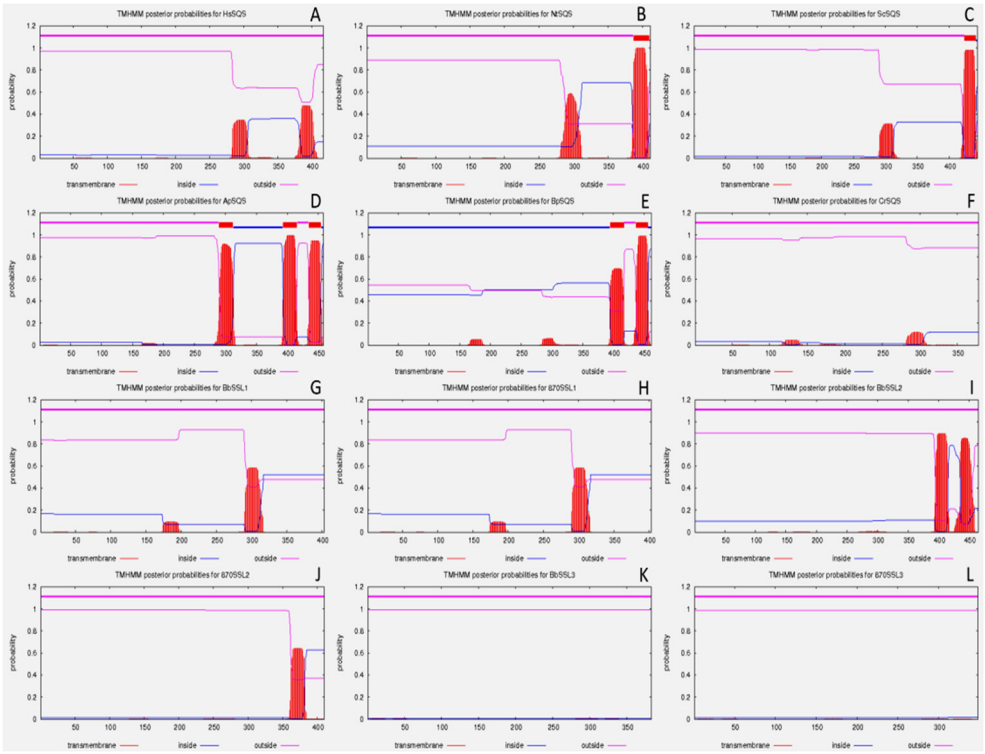

Figure 5. TMHMM plots generated for different SQS and SSL amino acids sequences. (A) *Homo sapiens* – HsSQS, (B) *Nicotiana tabacum* – NtSQS, (C) *Saccharamyces cerevisiae* –ScSQS, (D) *Auxenochlorella protothecoides* – ApSQS, (E) *Bathycoccus prasinos* – BpSQS, (F) *Chlamydomonas reinhardtii* – CrSQS, (G) *Botryococcus braunii* SSL-1 – BbSSL1, (H) AICB 870SSL-1 – 870SSL1, (I) *Botryococcus braunii* SSL-2 – BbSSL2, (J) AICB 870SSL-2 – 870SSL2, (K) *Botryococcus braunii* SSL-3 – BbSSL3, and (L) AICB 870SSL-3 – 870SSL3.

Only one TMM was observed in plants and yeast, in the following taxa: *Arabidopsis thaliana, Nicotiana tabacum* (Fig. 5 b)*, Solanum chacoense* and *Saccharomyces cerevisiae* (Fig. 5 c).

The highest rate of variability was observed in algae, the number of predicted TMM varies between three in *Auxenochlorella protothecoides* (Fig.5 d) and zero in the SQS from *Chlamydomonas reinhardtii* (Fig.5 f) and SSL from *Botryococcus* (Fig.5 g-l); as well, two THMM were observed in *Bathycoccus prasinos* (Fig.5 e).

Different papers about subcellular localization in rat hepatic cells (Stamellos *et al*., 1993) and yeasts (Zhang *et al*., 1993) showed that SQS is embedded in endoplasmic reticulum membrane through α-helix TMM. On the contrary, in this TMM prediction experiment we were unable to identify the membrane-spanning C-terminal motifs. Thus, more extensive investigation is further required.

Conclusions

In the present study we successfully identified and analyzed three squalene synthase-like cDNA fragments in the *Botryococcus terribilis* AICB870 stain.

Light and fluorescence microscopy observations showed that *Botryococcus* sp. AICB 870 presents similar features with a *Botryococcus terribilis* species, as described by Komárek and Marvan (1992), FanésTreviño *et al.* (2010), de Queiroz Mendes *et al.* (2012) and Hegedűs *et al.* (2014).

In order to extract high quantities of good quality RNA from *B. terribilis* AICB 870, the most suited method is that proposed by Ghawana *et al*. 2011. Using this protocol we were able to extract large quantity and high quality of RNA from the studied strain.

Based on the cDNA fragments from *B. braunii*, we were able to design new primers pairs, which successfully amplified the related cDNA fragments from *B. terribilis strain* AICB 870.

Bioinformatic analysis of nucleotide and translated amino acid sequences including G+C content, nucleotide frequencies, amino acids frequencies, computed Mw/pI and transmembrane motif prediction showed a high degree of similarity between the SSL identified as pertaining to *Botryococcus braunii* and those generated in the present work.

Multiple sequence alignments revealed four conserved domains, two aspartate-rich motifs and one NADPH binding domain including those from *B. terribilis* AICB 870, which were identified and analyzed for the first time in the present work.

Acknowledgements

This paper is a result of a doctoral research made possible by the financial support of the Sectoral Operational Programme for Human Resources Development 2007-2013, co-financed by the European Social Fund, under the project POSDRU/159/1.5/S/133391- "Doctoral and postdoctoral excellence programs for training highly qualified human resources for research in the fields of Life Sciences, Environment and Earth".

REFERENCES

Baba, M., Ioki, M., Nakajima, N., Shiraiwa, Y., Watanabe, M.M. (2012) Transcriptome analysis of an oil-rich race A strains of *Botryococcus braunii* (BOT-88-2) by *de novo* assembly of pyrosequencing cDNA reads, *Biores. Technol.,* **100**, 282-286

Bell, S.A., Niehaus, T.D., Nybo, S.E., Chappell, J. (2014) Structure - Function Mapping of Key Determinants for Hydrocarbon Biosynthesis by Squalene and Squalene Synthase-like Enzymes from the Green Alga Botryococcus braunii Race B, *Biochem*., **53**, 7570-7581

Blagg, B.S., Jarstfer, M.B., Rogers, D.H., Poulter, C.D. (2002) Recombinant squalene synthase. A mechanism for the rearrangement of presqualene diphosphate to squalene, *J. Am. Chem. Soc.*, **124** (30), 8846–8853

Brown, A.C., Knights, B.A., Conway, E. (1969) Hydrocarbon content and its relationship to physiological state in green alga *Botryococcus braunii*, *Phytochem.*, **8**, 543–547

de Queiroz Mendes, M.C., Comas González, A.A., Vieira Moreno, M.L., Pereira Figueira, C., de Castro Nunes, J.M. (2012) Morphological and ultrastructure features of a strain of *Botryococcus terribilis* (Trebouxiophyceae) from Brazil, *J. Phycol.*, **48,** 1099–1106

Devarenne, T.P., Shin, D.H., Back, K., Yin, S., Chappell, J. (1998) Molecular characterization of tobacco squalene synthase and regulation in response to fungal elicitor, *Arch. Biochem. Biophys.*, **349** (2), 205-215

Dragoș, N., Péterfi, L., Momeu, L., Popescu, C. (1997) An introduction to the algae and the culture collection of algae at the Institute of Biological Research Cluj-Napoca. Cluj University Press, Cluj-Napoca, Romania

FanésTreviño, I., Sánchez-Castillo, P., Comas González, A. (2009) Contribution to the taxonomic study of the family Botryococcaceae (Trebouxiophyceae, Chlorophyta) in southern Spain, *Cryptogamie Algol.,* **30**, 17–30

Farrell, R.E. (2010) RNA Methodologies – A Laboratory Guide for Isolation and Characterization, *4th Edition*, Academic Press - Elsevier, London, pp. 179-219

Gao, C., Wang, Y., Shen, Y., Yan, D., He, X., Dai, J., Wu, Q. (2014) Oil accumulation mechanisms of the oleaginous microalga *Chlorella protothecoides* revealed through its genome, transcriptomes, and proteomes, *BMC Genomics*, **15** (1), 582

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, in The Proteomics Protocols Handbook, Walker, J.M., (ed), Humana Press

Ghawana, S., Paul, A., Kumar, H., Kumar, A., Singh, H., Bhardwaj, P.K., Rani, A., Singh, R.S., Raizada, J., Singh, K., Kumar, S. (2011) An RNA isolation system for plant tissues rich in secondary metabolites, *BMC Research Notes*, **4** (85)

Ginzberg, I., Thippeswamy, M., Fogelman, E., Demirel, U., Mweetwa, A.M., Tokuhisa, J., Veilleux, R.E. (2012) Induction of potato steroidal glycoalkaloid biosynthetic pathway by overexpression of cDNA encoding primary metabolism HMG-CoA, *Planta*, **235** (6), 1341-1353

Gu, P., Ishii, Y., Spencer, T.A., 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

Guiry, M.D., and Guiry, G.M., (2015) AlgaeBase. World-wide electronic publication, National University of Ireland, Galway

Hegedűs, A., Mocan, A., Barbu-Tudoran, L., Coman, C., Drugă, B., Sicora, C., Dragoș, N. (2015) Morphological, biochemical, and phylogenetic assessments of eight Botryococcus terribilis strains collected from freshwaters of Transylvania, *J. App. Phycol.*, **27** (2), 865-878

Hegedűs, A., Coman, C., Drugă, D., Sicora, C., Dragoș, N. (2014) *Botryococcus terribilis* - A microalga capable to produce hydrocarbons similar to fossile fuel, *J. Biotechnol.,* **185**, S121-122

Hillen, L.W., Pollard, G., Wake, L.V., White, N. (1982) Hydrocracking of the oils of *Botryococcus braunii* to transport fuels, *Biotechnol. Bioeng.,* **24**, 193–205

Jennings, S.M., Tsay, Y.H., Fisch, T.M., Robinson, G.W. (1991) Molecular cloning and characterization of the yeast gene for squalene synthetase, *Proc. Natl. Acad. Sci. U.S.A.*, **88** (14), 6038-6042

Kalendar, R., Lee, D., Schulman, A.H. (2011) Java web tools for PCR, in silico PCR, and oligonucleotide assembly and analysis, *Genomics*, **98** (2), 137-144

Kawachi, M., Tanoi, T., Demura, M., Kaya, K., Watanabe, M.M. (2012) Relationship between hydrocarbon and molecular phylogeny of *Botryococcus braunii*, *Algal Res.,* **2**,114-119

Kim, B.-H., Ramanan, R., Cho, D.-H., Choi, G-G., La, H.-J., Ahn, C.-Y., Oh, H.-M., Kim, H.-S. (2012) Simple, rapid and cost-effective method for high quality nucleic acids extraction from different strains of *Botryococcus braunii*, *PLoS ONE*, **7** (5), e37770

Komárek, J., and Marvan, P. (1992) Morphological differences in natural populations of Botryococcus (Chlorophyceae), *Arch. Protistenkd.*, **141**, 65–100

Largeau, C., Casadevall, E., Berkaloff, C. (1980) The biosynthesis of the long-chain hydrocarbon in the green alga *Botryococcus braunii*, *Phytochem*., **19**, 1081–1085

Lee, S., and Poulter, C.D. (2008) Cloning, Solubilization, and Characterization of Squalene Synthase from *Thermosynechococcus elongatus* BP-1, *J. Bacteriol.*, **190** (11), 3808-3816

McKenzie, T.L., Jiang, G., Straubhaar, J.R., Conrad, D.G., Shechter, I. (1992) Molecular cloning, expression, and characterization of the cDNA for the rat hepatic squalene synthase, *J. Biol. Chem.*, **267** (30), 21368-21374

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., *et al*. (2007), The *Chlamydomonas* genome reveals the evolution of key animal and plant functions, *Science*, **318** (5848), 245-250

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

Metzger, P., David, M., Casadevall, E. (1987) Biosynthesis of triterpenoid hydrocarbons in the B-race of the green alga *Botryococcus braunii*. Sites of production and nature of the methylating agent, *Phytochem.*, **26**, 129–134

Metzger, P., Berkaloff, C., Casadevall, E., Coute, A. (1985) Alkadiene – and botryococcene - producing races of wild strains of *Botryococcus braunii*, *Phytochem.,* **24**, 2305–2312

Nakashima, T., Inoue, T., Oka, A., Nishino, T., Osumi, T., Hata, S. (1995) Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase, *Proc. Natl. Acad. Sci. U.S.A.*, **92** (6), 2328-2332

Niehaus, T.D., Okada, S., Devarenne, T.P., Watt, D.S., Sviripa, V., Chappell, J. (2011) Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii*, *Proc. Natl. Acad. Sci. U.S.A.*, **108** (30), 12260-12265

Okada, S., Devarenne, T.P., Murakami, M., Abe, H., and Chappell, J. (2004) Characterization of botryococcene synthase enzyme activity, a squalene synthase-like activity from the green microalga Botryococcus braunii, Race B, *Arch. Biochem. Biophys.,* **422**, 110−118

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

Pandit, J., Danley, D.E., Schulte, G.K., Mazzalupo, S., Pauly, T.A., Hayward, C.M., Hamanaka, E.S., Thompson, J.F., Harwood, Jr. H.J. (2000) Crystal structure of human squalene synthase. A key enzyme in cholesterol biosynthesis, *J. Biol. Chem.*, **275**, 30610–30617

Poulter, C.D. (1990) Biosynthesis of non-head-to-tail terpenes - formation of 1′-1 and 1′-3 linkages, *Acc. Chem. Res.*, **23**, 70–77

Sasiak, K., and Rilling, H.C. (1988) Purification to homogeneity and some properties of squalene synthetase, *Arch. Biochem. Biophys.*, **260**, 622–627

Sato, I., Ito, Y., Okada, S., Murakami, M., Abe, H. (2003) Biosynthesis of the triterpenoids, botryococcenes and tetramethyl squalene in the B race of *Botryococcus braunii* via the non-mevalonate pathway, *Tetrahedron Lett.,* **44**, 7035–7037

Schechter, I., Conrad, D.G., Hart, I., Berger, R.C., McKenzie, T.L., Bleskan, J., Patterson, D. (1994) Localization of the squalene synthase gene (FDFT1) to human chromosome 8p22-p23.1, *Genomics*, **20** (1), 116-118

Shechter, I., Klinger, E., Rucker, M.L., Engstrom, R.G., Spirito, J.A., Islam, M.A., Boettcher, B.R., Weinstein, D.B. (1992) Solubilization, purification, and characterization of a truncated form of rat hepatic squalene synthetase, *J. Biol. Chem.*, **267** (12), 8628-8635

Sonnhammer, E.L., von Heijne, G., Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences, *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **6**, 175-182

Stamellos, K.D., Shackelford, J.E., Shechter, I., Jiang, G., Conrad, D., Keller, G.A., Krisans, S.K. (1993) Subcellular localization of squalene synthase in rat hepatic cells. Biochemical and immunochemical evidence, *J. Biol. Chem.*, **268**, 12825–12836

Tamura, K., Strecher, G., Peterson, D., Filipski, A., Kumar, S. (2013) MEGA 6 Molecular Evolutionary Genetics Analysis Version 6.0, *Mol. Biol. Evol.*, **30** (12), 2725-2729

Weiss, T.L., Roth, R., Goodson, C., Vitha, S., Black, I., Azadi, P., Rusch, J., Holzenburg, A., Devarenne, T.P., Goodenough, U. (2012) Colony Organization in the Green Alga Botryococcus braunii (Race B) is Specified by a Complex Extracellular Matrix, *Euk. Cell*, **11** (12), 1424-1440

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction, *BMC Bioinformatics*, **13**, 134

Zhang, D., Jennings, S.M., Robinson, G.W., Poulter, C.D. (1993) Yeast squalene synthase: expression, purification, and characterization of soluble recombinant enzyme, *Arch. Biochem. Biophys.*, **304**, 133–143