

Modes of hydrocarbon oil biosynthesis revealed by comparative gene expression analysis for race A and race B strains of *Botryococcus braunii*

Motohide Ioki ^{a, 1}, Masato Baba ^{b, 1}, Haniyeh Bidadi ^{a, b}, Iwane Suzuki ^b, Yoshihiro Shiraiwa ^b, Makoto M. Watanabe ^b, Nobuyoshi Nakajima ^{a, b, *}

^a Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Ibaraki, 305-8506 Japan

^b Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

* Corresponding author: E-mail, naka-320@nies.go.jp; FAX, +81-29-850-2490

¹ These authors contributed equally to this work.

Abstract

To clarify the oil biosynthetic routes of the oil-producing green alga *Botryococcus braunii*, here the race-specific gene expression patterns were examined using representative strains of race A and race B producing fatty acid- and triterpene-derived hydrocarbon oils, respectively. The strain-specific gene expression patterns in the BOT-88-2 strain (race A) and the BOT-22 strain (race B) were revealed by transcriptome comparison and real-time PCR quantification. For race A, it was inferred from the gene expression patterns that the fatty acid elongation in the acyl-carrier-protein (acp)-bound form followed by further elongation in the coenzyme A (CoA)-bound form is the major route of oil biosynthesis. The fatty acids may be

desaturated in both acp- and CoA-bound forms and once metabolized into glycerolipids prior to further elongation. For race B, relatively direct entry of photosynthetic products from the reductive pentose phosphate cycle into the mevalonate-independent triterpene biosynthesis was implicated.

Keywords:

Alkadiene; Botryococcene; *Botryococcus braunii*; Gene expression; Hydrocarbon

1. Introduction

Botryococcus braunii produces large amounts of highly reduced long-chain liquid hydrocarbons resembling petroleum from photosynthetic products. Strains of *B. braunii* are classified into race A, B, or L according to the types of hydrocarbon oils they produce: very long-chain fatty acid (VLCFA)-, triterpene-, and tetraterpene-derived hydrocarbons, respectively. Race A and B strains are more abundant in nature and generally exhibit higher oil contents than race L strains (Reviewed in Banerjee et al., 2002; Metzger and Largeau, 2005). For improvement of the cost effectiveness of oil production using *B. braunii* for potential industrial use, understanding the oil biosynthesis pathways and their regulation is critical. To date, the hydrocarbon biosynthesis pathways in *B. braunii* have been inferred mainly from chemical structures of the hydrocarbons, feeding experiments using radioactive isotopes, and information about hydrocarbon biosynthesis in other organisms (Banerjee et al., 2002; Dennis and Kolattukudy, 1991). On the other hand, details of the oil biosynthetic routes and the oil biosynthesis genes are poorly understood. Majority of genes encoding enzymes

associated with hydrocarbon biosynthesis are unidentified with the exception of the squalene synthase gene and its homologs (Okada et al., 2000; Niehaus et al., 2011). To gain insights into the oil production in *B. braunii*, here transcriptome comparison and real-time PCR quantification were performed for determination of strain-specific expression patterns of associated genes in a representative race A strain (BOT-88-2) and a race B strain (BOT-22) exhibiting vigorous growth and oil production.

2. Materials and Methods

The EST datasets of BOT-88-2 and BOT-22 consisting of 29,038 and 27,427 non-redundant sequences (accession numbers FX056085 – FX085122 and FX085123 – FX112549) assembled from 183,250 and 206,261 filtered pyrosequencing reads, respectively, were used to compare the EST read counts of genes encoding hydrocarbon oil biosynthesis-related enzymes (Baba et al., 2012, this issue; Ioki et al., 2012a, this issue). Hydrocarbon biosynthesis-related genes were identified based on the enzyme nomenclature (accepted and alternative names) and the EST count for each candidate gene was determined using the EST Viewer software (Dragon Genomics Center, Takara Bio Co. Ltd, Japan) fed with BLASTX search hits from the non-redundant (nr) database (ver. 2009.01.09) compiled by National Center for Biotechnology Information (NCBI). The BLASTX search was performed with threshold bit score of 50, which corresponds to the *e*-value of approximately 1×10^{-5} . For comparison of EST counts between the two strains, the relative EST count for each gene was determined by multiplying, by 200,000, the proportion of the EST count of a specific gene in the total EST count for each strain.

Real-time PCR was carried out to examine the transcript accumulation levels for hydrocarbon biosynthesis-related genes in 4 biological replicates grown as described by Baba et al. (2012, this issue). For the enzymes whose candidate genes were found in the pyrosequencing EST datasets of BOT-88-2 or BOT-22, the non-redundant sequence exhibiting the largest EST counts was used to design the primers. Additional sequences from the BOT-70 dataset were used for the botryococcene biosynthesis-related genes missing in the pyrosequencing datasets. Throughout the experiments, the α -tubulin expression level was used as the standard for normalization because the EST counts were comparable in the two strains. For each gene, the primers were optimized based on the melting curve of the PCR product as well as the PCR amplification efficiencies calculated using dilution series of cDNA or genomic DNA. Genomic DNA was used as a template in the efficiency calculation for weakly expressed genes for which cDNA could not be used as a template for reliable efficiency calculation. The PCR amplification efficiency was also used for data normalization. The cDNA was synthesized from total RNA (100 ng/ reaction) with an oligo(dT) primer using First Strand cDNA Synthesis Kit (Roche Applied Science, Eugene, OR, USA) and used as a template in the real-time PCR reaction. The real-time PCR reactions were performed using LightCycler 480 SYBR Green I Master (Roche Applied Science). The mRNA levels were calculated via the delta-delta CT method using LightCycler 480 SYBR Software Version 1.5 (Roche Applied Science).

3. Results and Discussion

Strain-specific expression profiles of hydrocarbon biosynthesis-related genes

were disclosed by EST count comparison (Table 1, Table 2, Table S1, Table S2) and real-time PCR quantification (Table 3, Fig. 1). The EST count comparison using approximately 200,000 cDNA reads for each strain allowed global comparison between the two strains, but only one EST dataset was available for each strain (Baba et al., 2012, this issue; Ioki et al., 2012a, this issue). Therefore, real-time PCR quantification was performed to examine the gene expression patterns in multiple biological replicates. The expression levels of hydrocarbon genes critically differed between BOT-88-2 and BOT-22, providing inferences on the race A and race B oil biosynthesis pathways and potential targets of genetic engineering for improved oil productivity.

The race A hydrocarbon products, alkadienes and alkatrienes, are believed to be biosynthesized from fatty acids. It is still unclear whether the fatty acid elongation takes place in the acyl-carrier-protein (acp)- or coenzyme A (CoA)-bound form in *B. braunii*. To determine what synthetic routes are active in the race A oil biosynthesis, gene expression patterns for fatty acid biosynthesis were compared between a race A strain (BOT-88-2) and a race B strain (BOT-22). The transcriptomes of two *B. braunii* races indicated that two genes related to the acyl-acp elongation, acetyl-CoA carboxylase (EC 6.4.1.2) and 3-oxoacyl-[acyl-carrier-protein] synthase I (EC 2.3.1.41), were vigorously expressed only in BOT-88-2 (Table 1, Supplementary Table S1). The real-time PCR quantification also showed that expression levels of these two genes of BOT-88-2 were much higher (5 to 189,000-fold higher) than those of BOT-22 (Fig. 1a). In contrast, we detected no ESTs for acetyl-CoA C-acetyltransferase (EC 2.3.1.9) and acetyl-CoA C-acyltransferase (EC 2.3.1.16) that were essential for the first step of acyl-CoA elongation pathway (Table 1, Supplementary Table S1). These data suggest that the acyl-acp elongation pathway is predominant in the race A oil biosynthesis.

The second question regarding the race A oil biosynthesis was whether the fatty acid desaturation takes place in the form of acyl-*acp* or acyl-CoA. The real-time quantification indicated that expression levels of both stearoyl-CoA 9-desaturase (EC. 1.14.19.1) and acyl-*acp* desaturase (EC. 1.14.19.2) were 43,300 to 15,100-fold higher in BOT-88-2 than in BOT-22, suggesting the involvement of both types of enzymes in the race A oil biosynthesis (Fig. 1a). Fatty acid desaturation in the CoA-bound form is uncommon in photosynthetic organisms (Los and Murata, 1998).

The third question was whether the unsaturated fatty acids directly enter the VLCFA elongation pathway or once metabolized into glycerolipids. Both the EST count comparison and real-time PCR quantification indicated strikingly strong and strain-specific expression of genes associated with glycerolipid metabolism in BOT-88-2 (Table 1, Fig. 1b). For majority of glycerolipid metabolism-related genes, ESTs were retrieved only from the BOT-88-2 dataset (Table 1, Supplementary Table S1). Interestingly, the gene expression for the entry point into the VLCFA elongation pathway catalyzed by 3-ketoacyl-CoA synthase (EC 2.3.1.-) was 48,000,000-fold lower in BOT-88-2 than in BOT-22 (Fig. 1a). These results suggested that the substrate entry into the VLCFA elongation pathway might be regulated through glycerolipid metabolisms. The entry point into the VLCFA elongation pathway seemed to be the rate-limiting step and is a potential target of genetic engineering for improved oil productivity in the race A strain.

The race B hydrocarbon products, botryococcenes, are synthesized from intermediates called isopentenyl-pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). The mevalonate pathway and the mevalonate-independent pathway are known for the biosynthesis of IPP and DMAPP. Synthesis of IPP and DMAPP is

followed by triterpene synthesis for botryococcene formation. Gene expression patterns indicated that the mevalonate-independent pathway is the major route of the race B oil biosynthesis. Candidate genes for enzymes catalyzing 10 out of 11 reactions in botryococcene biosynthesis via the mevalonate-independent pathway have been discovered from the BOT-88-2, BOT-22, or BOT-70 EST datasets (Ioki et al., 2011a, b, this issue). The only gene missing is the gene for 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (EC 2.7.7.60) (Table 2, Supplementary Table S2).

The race B oil biosynthesis via the mevalonate-independent pathway begins with either formation of 1-deoxy-D-xylulose-5-phosphate or xylulose-5-phosphate. Pyruvate and glyceraldehyde-3-phosphate are the starting substrates for the conventional entry route via 1-deoxy-D-xylulose-5-phosphate (Rohmer, 1999). Interestingly, this conventional entry route into the mevalonate-independent pathway has been reported to be inactive in a cyanobacterium *Synechocystis* sp. strain PCC6803 under photosynthetic conditions. In the cyanobacterium, photosynthetic products from the reductive pentose phosphate cycle (or the Calvin-Benson cycle) such as xylulose-5-phosphate are the starting substrates for the IPP synthesis (Ershov et al., 2002; Poliquin et al., 2004). It has not been clear whether the conventional or the alternative entry route into the mevalonate-independent pathway is active in *B. braunii*. The expression level determined via real-time quantification was 7-fold lower in BOT-22 than in BOT-88-2 for the gene putatively encoding 1-deoxy-D-xylulose-5-phosphate synthase (EC 2.2.1.7), indicating that the conventional entry route into the mevalonate-independent pathway is not highly active in *B. braunii*. The alternative entry route into the mevalonate-independent pathway seemed active in

BOT-22 because of the strong expression of transketolase (EC 2.2.1.1) and transaldolase (EC 2.2.1.2). Genes for the subsequent reactions catalyzed by (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (EC 1.17.7.1) and 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2) were expressed in BOT-22 at levels more than 1,000 times higher than BOT-88-2 (Fig. 1c).

All of the genes putatively encoding the enzymes of the triterpene synthesis pathway were detected in both BOT-88-2 and BOT-22 via real-time PCR. No statistically significant differences were detected for the initial steps of the triterpene synthesis pathway catalyzed by isopentenyl-diphosphate Delta-isomerase (EC 5.3.3.2) and dimethylallyltranstransferase (EC 2.5.1.1). Among the triterpene synthesis-related genes, statistically significant differences between the BOT-88-2 and BOT-22 were detected only for geranyltranstransferase (EC 2.5.1.10) and squalene synthase (EC 2.5.1.21). For these genes, the expression levels were 5,6000 to 1,910,000 times higher in the BOT-22 strain, indicating association with the race B oil biosynthesis (Fig. 1c).

The EST retrieval and real-time PCR quantification lead to inferences on the regulation of the botryococcene biosynthesis. Initial reactions of the mevalonate-independent pathway, namely, the reaction catalyzed by 1-deoxy-D-xylulose-5-phosphate synthase (EC 2.2.1.7), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (1.1.1.267), and 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (EC 2.7.7.60), seemed to be the rate-limiting steps of the race B oil biosynthesis because the expression levels were not necessarily high despite its association with the race B oil biosynthesis (Table 2, Table S2, Fig. 1c). Initial reactions of the triterpene synthesis pathway, namely, the steps catalyzed by isopentenyl-diphosphate Delta-isomerase (EC 5.3.3.2) and

dimethylallyltranstransferase (EC 2.5.1.1), also seemed to be the rate-limiting steps because they were not strongly expressed in BOT-22 despite their involvement in the race B oil biosynthesis (Table 2, Fig. 1c). These weakly expressed genes are potential targets of genetic engineering for improved oil productivity in the race B strain.

4. Conclusions

Analysis of race-specific gene expression patterns provided insights into details of hydrocarbon oil biosynthesis in *B. braunii*. Potential targets of enzymatic steps for improved oil productivity proposed based on the race-specific gene expression patterns in this paper are expected to accelerate the urgent research and development for commercialization of the algal oil.

Acknowledgements

This research was supported by the Core Research for Evolutionary Science and Technology program of Japan Science and Technology Agency.

References

Baba, M., Ioki, M., Nakajima, N., Shiraiwa, Y., Watanabe, M.M., 2012. Transcriptome analysis of an oil-rich race A strain of *Botryococcus braunii* (BOT-88-2) by *de novo* assembly of pyrosequencing cDNA reads. *Bioresour Technol.* 109, 282-286.

- Banerjee, A., Sharma, R., Chisti, Y., Banerjee, U.C., 2002. *Botryococcus braunii*: a renewable source of hydrocarbons and other chemicals. *Crit Rev Biotechnol.* 22(3), 245-279.
- Dennis, M.W., Kolattukudy, P.E., 1991. Alkane biosynthesis by decarbonylation of aldehyde catalyzed by a microsomal preparation from *Botryococcus braunii*. *Arch. Biochem Biophys.* 287(2), 268-275.
- Ershov, Y.V., Gantt, R.R., Cunningham, F.X. Jr., Gantt E., 2002. Isoprenoid biosynthesis in *Synechocystis* sp. Strain PCC6803 is stimulated by compounds of the pentose phosphate cycle but not by pyruvate or deoxyxylulose-5-phosphate. *J. Bacteriol.* 184(18), 5045-5051.
- Ioki, M., Baba, M., Nakajima, N., Shiraiwa, Y., Watanabe, M.M., 2012a. Transcriptome analysis of an oil-rich race B strain of *Botryococcus braunii* (BOT-22) by *de novo* assembly of pyrosequencing cDNA reads. *Bioresour Technol.* 109, 292-296.
- Ioki, M., Baba, M., Nakajima, N., Shiraiwa, Y., Watanabe, M.M., 2012b. Transcriptome analysis of an oil-rich race B strain of *Botryococcus braunii* (BOT-70) by *de novo* assembly of 5'-end sequences of full-length cDNA clones. *Bioresour Technol.* 109, 277-281.
- Los, D.A., Murata, N., 1998. Structure and expression of fatty acid desaturases. *Biochim Biophys Acta.* 1394, 3-15.

Metzger, P., Largeau, C., 2005. *Botryococcus braunii*: a rich source for hydrocarbons and related ether lipids. *Appl Microbiol Biotechnol.* 66(5), 486-496.

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., Chappell, J., 2000. Molecular characterization of squalene synthase from the green microalga *Botryococcus braunii*, race B. *Arch. Biochem. Biophys.* 373(2), 307-317.

Poliquin, K., Ershov, Y.V., Cunningham, F.X. Jr., Woreta, T.T., Gantt, R.R., Gantt, E., 2004. Inactivation of *sll1556* in *Synechocystis* strain PCC 6803 impairs isoprenoid biosynthesis from pentose phosphate cycle substrates in vitro. *J. Bacteriol.* 186(14), 4685-4693.

Rohmer, M., 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat Prod Rep.* 16, 565–574.

Figure captions

Fig. 1 Quantification of mRNA levels for genes associated with hydrocarbon oil biosynthesis in BOT-88-2 and BOT-22. Transcript accumulation levels were

determined via real-time PCR for genes associated with unsaturated VLCFA biosynthesis **(a)**, glycerolipid metabolism **(b)**, and botryococcene biosynthesis **(c)**. The mRNA accumulation levels were examined in 4 biological replicates each of BOT-88-2 and BOT-22. The fold difference in mRNA abundance was calculated by dividing the level in BOT-88-2 by that in BOT-22 for each gene except for the 1-deoxy-D-xylulose-5-phosphate reductoisomerase (1.1.1.267) gene, whose transcript was not detected in BOT-22. Asterisks indicate statistical differences between the two strains determined by *t*-test at $P < 0.1$ (*) and $P < 0.01$ (**).

P.S. Please check Fig. 1 and Tables at

<http://dx.doi.org/10.1016/j.biortech.2011.11.078>