PHYLOGENETICAL APPROACH FOR THE SEARCH OF VALUABLE METABOLIC PRODUCTS IN CYANOBACTERIA

Liliana Cepoi[1], Jacob Golan[2], Andrii P. Gryganskyi[3]

[1]*Institute of Microbiology and Biotechnology of the ASM, R. Moldova*
[2]*University of Wisconsin-Madison, Madison, Wisconsin, USA*
[3]*DukeUniversity, Durham, NC, USA*

Introduction

Cyanobacteria are an important group of microorganisms that play crucial role in both terrestrial and aquatic ecosystems. Species are beneficial as a food source, as well as producers of carotenoids, antioxidants and other secondary metabolites.

A phylogenetic approach may be an indispensable tool to detect strains of potential value for the production of various secondary metabolites and to insure the quality and safety of biological products. However, most existing phylogenetic studies lack precision and use mostly rDNA or single gene approaches, thus preventing a more detailed understanding of the cyanobacteria evolutionary trajectories.

Advances in molecular and genome studies have revealed the main genes involved in different stages of secondary metabolites production. Several genes have been identified de novo in triacylglycerol (TAG) biosynthesis pathway and are known to be involved in lipid synthesis (*rbsL, me g6562, accA, accD, dgat g2354, dgat g3280, gat g7063*), production of malic enzyme subunits (ACCase and diacylglycerol acyl transferase), and carotenoid synthesis (*crtB, crtP, crtQ-1, crtQ-2, crtL, crtL-2, crtO, crtR*). The first two gene groups are well studied in relation to biofuel production, while the third group has been widely employed in the pharmaceutical, cosmetics, and food industry for natural dyes and antioxidants.

Carotenoids are a specific metabolite involved in the photosynthesis performed by cyanobacteria. The carotenoids serve as light-harvesting pigments in photosynthesis, and protect the cells against photooxidative damage [4]. Carotenoids play a central role in the deactivation of 3Chl* and 1O2*, and the reduction of reactive oxygen species (ROS) formation due to the thermal dissipation of excess light energy at the level of 1Chl* [6]. In addition carotenoids are able to mitigate the effects of ROS such as superoxide and peroxilipids radicals [3, 11, 14].

In cyanobacteria, various carotenoids have been identified, including the unique ketocarotenoids, echinenone and 4-ketomyxol, and the carotenoid glycosides, myxol glycosides and oscilloldiglycosides (β-carotene, β-criptoxanthin, zeaxanthin, caloxanthin, lutein, lycopene, nostoxanthin, echinenone, canthaxanthin, 3-OH echinenone, deoxymyxol, myxol, ketomyxol, hidroxymyxol, oscillol etc.). Major carotenoids in cyanobacteria include b-carotene, its hydroxyl and keto derivatives, and the carotenoid glycosides [16]. It is well known that the carotenoid composition depends on growth conditions such as light intensity, growth stage, composition of nutrient medium, and intensity of ROS formation [5, 9, 10]. However, the diversity in carotenoid composition might be due to the presence of specific carotenogenesis pathways.

In the biosynthesis pathway of carotenoids, four enzymes convert geranylgeranyl pyrophosphate to *β*-carotene: phytoene synthase, phytoenedesaturase, ζ-carotene desaturase and lycopene cyclase. Four desaturation steps are needed in the conversion from phytoene to lycopene, and two distinct pathways are known among cyanobacteria: the plant type and the bacterial type [16]. The plant-type requires three enzymes: *CrtP* (phytoenedesaturase), *CrtQ* (z-carotene desaturase) and *CrtH* (*cis*-carotene isomerase), and the bacterial type uses only one enzyme, *CrtI* (phytoenedesaturase) to convert phytoene to lycopene. The bacterial type is rarely found in cyanobacteria, and only in strains which retain ancestral properties of carotenoid biosynthesis. These observations

suggest that in the evolution of cyanobacteria species, the gene *CrtI* was replaced by *CrtP* [19]; and that the gene *CrtP* is found throughout species of cyanobacteria, playing an important role in the synthesis of carotenoids. For these reasons, a phylogenetic approach can be used to detect potentially efficient carotenoid producers, as well as identify unknown isolates and suitable primers for screening for useful genes.

Materials and methods

Genes *CrtP* of 14 cyanobacteria were found using BLAST search algorithm [1] of *Anabaena sp.* PCC 7210 *CrtP* from NCBI Nucleotide database. For all samples (excluded uncultured bacterium and identical *Anabaena = Nostoc* PCC 7120 samples) corresponding 16S sequences were found (Table 1). Resulted sequences (15 for *crtP1* and 13 for rRNA small 16S subunit) were combined in correspondent single text files (format fasta) and aligned using MUSCLE 3.5 [2] in Mesquite 3.03 [8]. Alignments were examined visually and ambiguous regions were adjusted manually resulting in 1350 characters from 1910 base pairs alignment total length for *crtP1* and 1479 vs. 1489 nucleotide base pairs for rRNA 16S. Maximum Likelihood (ML) approach was applied to generate phylogenetic trees using GARLI 1.1 [20]. We run 5000000 iterations using GTRGAMMA rates model. Phylogenetic support was assessed by 100 replication of bootstrap analysis using PAUP* 4.0a109 [13].Conserved regions in *CrtP* alignment were found visually based on identity of first and second codon positions with no or some differences (less than in five sequences from 15) in third codon positions.

Table 1. Sequences of *crtP1* gene and rDNA small 16S subunit used for phylogenetic analysis

Species	Strain	*CrtP*		16S rRNA	
		Accession #/ position in genome	Product/ Protein ID	NCBI Reference Sequence	Product
Anabaena sp.	PCC 7120	Y15114	phytoene desaturase CAB56040.1	See *Nostoc* sp.PCC 7120	*Ibid.*
Anabaena cylindrica	PCC 7122	CP003659 Reg. 4412834.. 4414607	phytoene desaturase AFZ59250.1	KM019919.1	16S ribosomal RNA
Anabaena variabilis	ATCC 29413	CP000117 Reg.: 6025392.. 6027184	phytoene desaturase BAB73531.1	NR_074300.1	16S ribosomal RNA
Calothrix sp.	PCC 7507	CP003943 Reg.: 4124330.. 4125464	phytoene phytoene desaturase CAB 56040.1 desaturase AFY 34025.1	NR_102891.1	16S ribosomal RNA
Cylindrospermum stagnale	PCC 7417	CP003642 Reg.: 3596561.. 3597711	phytoene desaturase AFZ25297.1	NR_102462	16S ribosomal RNA
Geitlerinema sp.	PCC 7407	CP003591 Reg.: 1360418.. 1361866	phytoene desaturase AFY65636.1	NR_102448.1	16S ribosomal RNA
Nodularia spumigena	CCY 9414	CP007203 Reg.: 2429160.. 2430283	phytoene desaturase AHJ28780.1	CP007203.2 Reg 1454563… 1456044	16S ribosomal RNA

‚Nostoc azollae'	0708	CP002059 Reg.: 2958611.. 2960379	phytoene desaturase ADI64708.1	NR_074259.1	16S ribosomal RNA
Nostoc punctiforme	PCC 73102	CP001037 Reg.: 3436797.. 3438236	Amine oxidase ACC81307.1	NR_074317.1	16S ribosomal RNA
Nostoc sp.	PCC 7107	CP003548 Reg.: 4841724.. 4843495	phytoene desaturase AFY44694.1	CP003548.1 Reg. 201748… 203226.	16S ribosomal RNA
Nostoc sp.	PCC 7120	BA000019 REG.: 2196420.. 2198212	phytoene desaturase BAB73532.1	NR_074310.1	16S ribosomal RNA
Nostoc sp.	PCC 7524	CP003552 Reg.: 5638496.. 5640281	phytoene desaturase AFY50457.1	CP003552.1 Reg. 2137997… 2139363	16S ribosomal RNA
Oscillatoria acuminata	PCC 6304	CP003607 Reg.: 1499581.. 1500940	phytoene desaturase AFY80941.1	NR_102463.1	16S ribosomal RNA
Rivularia sp.	PCC 7116	CP003549 Reg.: 5251957.. 5253652	phytoene desaturase AFY56622.1	NR_102458.1	16S ribosomal RNA
Uncultured bacterium	clone 66415	KP445989 Reg.: 5912..7700	phytoene desaturase	n. a.	n. a.

n. a. – not applicable

Results and discussions

We conducted an ML analysis of gene *crtP1* on 15 cyanobacteria samples in order to reveal phylogenetic relationships between these taxa. Closely related taxa in which the *CrtP* gene had not been previously described, should contain copies of the gene that vary only slightly. An ML analysis of the gene *CrtP* in 15 taxa of cyanobacteria has shown that the placement of species on a phylogenetic tree does not necessarily correspond to their taxonomical affiliation (Fig. 1). Although *Anabaena* group (three taxa) is basal for the whole tree, it includes one of the *Nostoc* samples (PCC7120) and does not have any significant statistical support. Other Noctoc *CrtP* genes are dispersed all over the tree and do not build any stable clade. Additionally, Nostoc *CrtP* genes are located together with any other gene isolated from this genus. This tree topology might be a result of misidentification, but more likely it reflects some problems in the current taxonomical boundaries of the genus *Nostoc*. We observed similar patterns for *Anabaena* samples (Fig. 1). These two cyanobacteria groups might be artificial and polyphyletic, supported by 16S and *hetR* single gene phylogenies [18], and a genome-wide cyanobacteria phylogeny [12]. Another explanation for the scattered placement of *CrtP* in taxonomically closely related taxa might be horizontal gene transfer. However, more isolates and genes should be included in such an analysis to detect possible horizontal gene transfer events in the evolution of branching filamentous cyanobacteria.

In addition, the topology of *CrtP* and 16S trees is not symmetrical (Fig. 1-2). Similarly, on both trees, *Geitlerinema* and *Oscillatoria* are nested together and this clade has significant bootstrap support. In general, *CrtP* tree has more statistical support compared to rDNA 16S tree, mostly because of more conserved and similar sequences, which produced better alignment. For other taxa that follow some similarity in topology amongst separate clades, sample placement impedes phylogenetic resolution. This might be due to different evolutionary trajectories of *CrtP* and the small ribosomal

subunit. Our trees are also difficult to compare with other phylogenetic schemes of branching filamentous cyanobacteria [Tomitany 2007; Shih et al. 2012] mostly due to sampling quantity. Nonetheless, a common feature to all trees discussed here is wide scale polyphily of various taxa, as shown in data sets of *Anabaena* and *Nostoc* (Fig. 1-2).

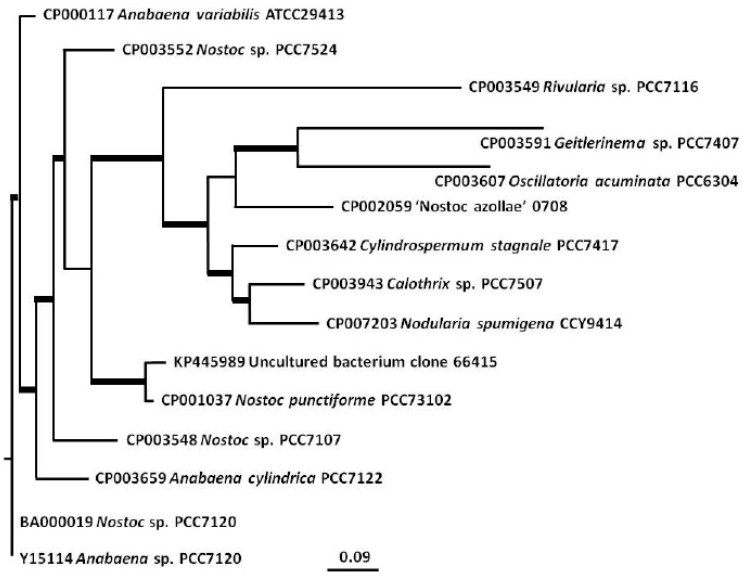

Figure 1. ML tree of crtP1 gene for 15 cyanobacteria isolates. Thick lines show the branches with sufficient (over 70%) bootstrap support.

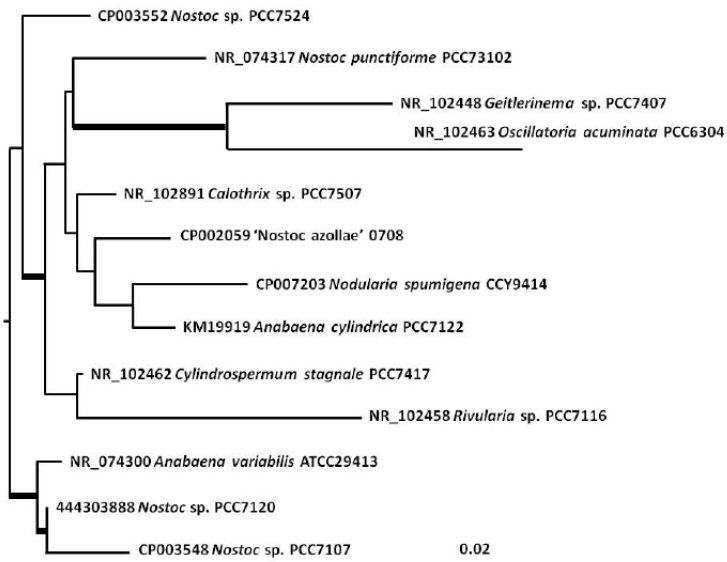

Figure 2. ML tree of 16S rRNA gene for 13 cyanobacteria isolates. Thick lines show the branches with sufficient (over 70%) bootstrap support.

Phylogenetic analysis can be helpful in the identification of unknown isolates. Thus on our tree, the *CrtP* gene of an unknown bacterium (clone 66415, acc. # KP445989)

is nested together with the gene of *Nostoc punctiforme.* These sequences possibly represent the same or closely related species - for example, the sequences of *Anabaena sp.* PCC 7120 (acc. # Y15114) and *Nostoc sp.* PCC 7120 (acc. # BA000019) are identical as Kaneko et al. (2001) suggest; they occupy the same position on the tree [7].

Alignment of amino acid sequences of the genes can also be used for primer designed to screen useful genes for DNA amplification. For the *CrtP* gene we have determined the fragment with total length 1350 nucleotide base pairs (450 amino acids) with no introns and a uniform structure, excluding *Anabaena cylindrica* (strain PCC 7122, genome acc. # CP003659) which contains a small insert of seven amino acids in the beginning of the sequence following the first conserved region. Both starting and ending part of this fragment with length of 39 nucleotide base pairs are well conserved (same amino acids with no changes or minor differences in 3rdcodon positions) are very suitable for primer design. The total length of the fragment allows significant overlap for its amplification using forward and reverse Sanger sequencing (200-300 bp). This fragment contains seven other well conserved regions each with a length over 20 nucleotides that is needed to cover the sufficient primer length (Fig. 3).

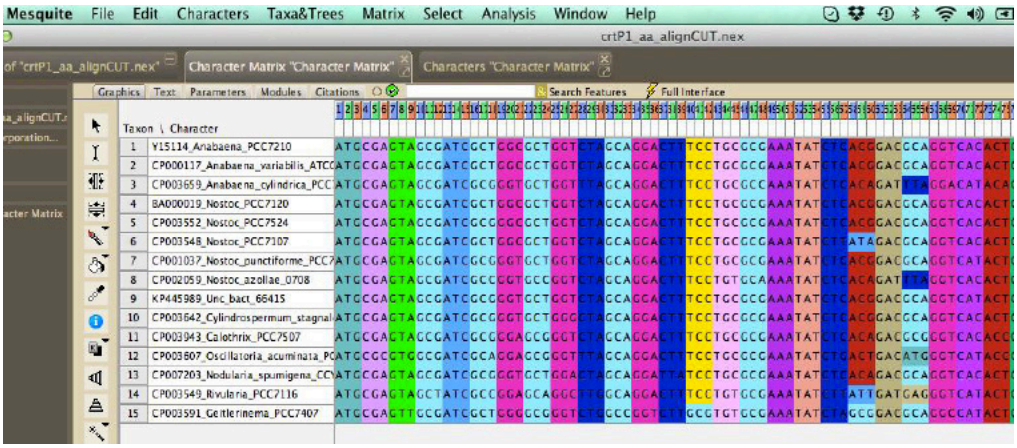

Fig. 3. Alignment of crtP1 gene with starting region (positions 1- 39) quite suitable for primer design.

Thus, phylogenetic analysis of single gene sequences of our selected data set can be useful in three instances: 1) targeting of prospective candidates of useful metabolites through revealing the phylogenetic relationships between them, 2) identification of unknown isolates, 3) screening of targeted genes using PCR methods.

Bibliography

1. Altschul S.F. et al. "Basiclocalalignmentsearchtool."// J. Mol. Biol. 2009, vol. 215, 403-410.

2. Edgar R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput.// Nucleic Acids Research. 2004, vol. 32(5), 1792-97.

3. Goiris K. et al. Antioxidant potential of microalgae in relation to their phenolic and carotenoid content.// J Appl Phycol. 2012, vol. 24, 1477–1486.

4. Hirschberg J., Chamovitz D. Catotenoids in Cyanobacteria.// The Molecular Biology of Cyanobacteria, Advances in Photosynthesis and Respiration. 2004, vol. 1, 559-579.

5. Iaţco Iu. Tehnologii de obţinere a preparatelor lipidice din biomasa algei verzi Dunaliella salina.// Teză de dr. în biol. Chişinău, 2012, 168 p.

6. Jahns P., Holzwarth A.R. The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II.// Biochim Biophys Acta. 2012, vol. 18(17), 182–193.

7. Kaneko T. et al. Complete genomic sequence of the filamentous nitrogen-fixing cyano-

bacterium Anabaena sp. strain PCC 7120.// JOURNAL DNA Res. 2001, vol. 8 (5), 205-213.

8. Maddison W. P. and Maddison D. R. Mesquite: a modular system for evolutionary analysis.// 2015, Version 3.03 http://mesquiteproject.org.

9. Miscu V. Biotehnologii de obținere a preparatelor pe bază de astaxantină din Haematococcus pluvialis flotow CNW-AV-05.// Teză de dr. în biol. Chișinău, 2010, 158p.

10. Özge Sözer. Carotenoids assist in the assembly and the functions of photosynthetic complexes in cyanobacteria.// Summary of Ph.D. thesis. Szeged, 2011, 11 p.

11. Ramel F. et al. Chemical quenching of singlet oxygen by carotenoids in plants.// Plant Physiol. 2012, vol. 158, 1267–1278.

12. Shih PM, et al. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing.// Proc Natl Acad Sci U S A. 2013, vol. 110(3), 1053-8. doi: 10.1073/pnas.1217107110.

13. Swofford DL. Phylogenetic Analysis Using Parsimony (*and Other Methods).// Version 4. Sunderland, Massachusetts: Sinauer Associates. 2004.

14. Takaichi S. Carotenoids in algae: distributions, biosyntheses andfunctions.// Mar Drugs. 2011, vol.9, 1101–1118.

15. Takaichi S., Mochimaru M. and Maoka T. Presence of free myxol and 4-hydroxymyxol and absence of myxol glycosides in Anabaena variabilis ATCC 29413, and proposal of a biosynthetic pathway of carotenoids.// Plant Cell Physiol. 2006, vol. 47, 211 – 216.

16. Takaichi S., Mochimaru M. Carotenoids and carotenogenesis in cyanobacteria: unique ketocarotenoids and carotenoid glycosides.// Cell. Mol. Life Sci. 2007, vol. 64, 2607 – 2619.

17. Takaichi S., Mochimaru M., Maoka T. and Katoh H. Myxol and 4-ketomyxol 2'-fucosides, not rhamnosides, from Anabaena sp. PCC 7120 and Nostoc punctiforme PCC 73102, and proposal for the biosynthetic pathway of carotenoids.// Plant Cell Physiol. 2005, vol. 46, 497 – 504.

18. Tomitani A. Evolution of branching filamentous cyanobacteria: Molecular-phylogenetic analyses of stigonematalean species.// Frontier Research on Earth Evolution, vol. 2, Institute for Frontier Research on Earth Evolution, Yokosuka, Japan, 2007, p 1-5.

19. TsuchiyaT. et al. The cyanobacterium Gloeobacter violaceus PCC 7421 uses bacterial-type phytoene desaturase in carotenoid biosynthesis.// FEBS Lett. 2005, vol. 579, 2125 – 2129.

20. Zwickl D. J. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion.// Ph.D. dissertation, The University of Texas at Austin, 2006.