J. Indian bot. Soc. (2024) Vol. 104 (4): 234-238

e-ISSN:2455-7218, ISSN:0019 - 4468

https://doi: 10.61289/jibs2024.09.30.1188

www.indianbotsoc.org

RESEARCH ARTICLE

A Comparative Study of the Rubisco Small Subunit Gene Sequence in a Few Taxa

Sachita P. Meshram, Rupesh S. Badere

Abstract

Fifty-one taxa were investigated to understand the variation in the gene sequence of the small subunit of Rubisco (rbcS). The nucleotide sequences of rbcS gene from GenBank were downloaded to generate a multiple sequence alignment using BioEdit. Later, a dendrogram was drawn from this alignment using Mesquite and the trees were searched using Nearest neighbour interchange and Subtree pruning & regrafting rearrangers. The resulting dendrograms resembled in topology with respect to the arrangement of a few taxa, while they differed with respect to the others. The present investigation reveals the considerable diversity in the nucleotide sequence of the rbcS gene among the taxa studied.

Keywords: Bioedit, dendrogram, clade, Mesquite, MSA.

Introduction

Ribulose bisphosphate carboxylase oxygenase or RuBisCo is an important enzyme for carbon fixation in photoautotrophs like photosynthetic bacteria, algae and plants. The enzyme mediates the acceptance of atmospheric CO_2 by ribulose 1,5-bisphosphate, which subsequently splits into two molecules of 3-phosphoglycerates (3PGA). This 3PGA later is used to synthesize carbohydrate and other organic compounds by the organisms. The primitive Rubisco was a homodimer of large subunits, which through several evolutionary steps, got the present-day form of a large complex protein. The modern Rubisco is organised in a higher oligomeric state composed of large and small subunits. This structural change in the enzyme has increased its catalytic efficiency and specificity (Erb and Zarzycki 2018).

Department of Botany, RTM Nagpur University, MJP Educational Campus, Amravati Road, Nagpur

***Corresponding Author:** Rupesh S. Badere, Department of Botany, RTM Nagpur University, MJP Educational Campus, Amravati Road, Nagpur- 440033, India, Email: rsbadere@nagpuruniversity.nic.in

How to cite this article: Meshram, S.P., Badere, R.S. (2024). A Comparative Study of the Rubisco Small Subunit Gene Sequence in a Few Taxa. *J. Indian bot. Soc.*, 104(4):234-238 Doi: 10.61289/jibs2024.09.30.1188

Source of support: Nil

Conflict of interest: None.

The small subunit (SSU) of the Rubisco is encoded by the rbcS gene (Jarvis and Soll 2001) belonging to a multigene family and is present in the nucleus (Suzuki 2009). The evolution has mainly altered the exon structure of SSU gene leading to a substantial modification in the encoded protein (Goldschmidt-Clermont and Rahire 1986). Similarly, the additional motif sequences have been found at the transit peptide region of rbcS gene in plants (Razzak 2017). The molecular phylogenetic studies discovered evolutionary diversification of SSU in the βA-βB loop region which influences the catalytic performance of the enzyme (Xiang et al. 2016).

A comparative study of the sequence of the *rbcS* in various photosynthetic organisms would aid in understanding the evolutionary tendency of the SSU subunit of the enzyme. Therefore, the present study was undertaken to establish the relationship among the *rbcS* sequences from organisms of diverse lineages.

Material and Methods

Drawing the Dendrogram

Nucleotide database of NCBI (available at http://www/ncbi/nlm.nhi.gov/) was searched for the cDNA sequences of *rbcS* gene among several prokaryotes and eukaryotes. The sequences were imported in a sequence alignment editor software – BioEdit (Hall 1999). Initially, the multiple sequence alignment (MSA) was generated by executing the relevant

Received: 13/03/2024 **Accepted:** 15/05/2024 **Published online**: 30/09/2024

commands in Bioedit and saved in the FASTA format. Prior to this the conserved regions in the MSA were also searched. Subsequently, dendrogram showing relationship among the taxa was drawn using Mesquite ver. 3.4.0 (Maddison and Maddison 2017). The MSA saved in FASTA format was opened in Mesquite and a heuristic search for 100 trees was made based on the character matrix of the MSA. Based on the rearrangement of subtrees, two types of rearrangers viz., the Nearest neighbour interchange rearranger (NNI rearranger) and the Subtree pruning & grafting rearranger (SPR rearranger) were used to generate the trees. Subsequently, a consensus tree was drawn for the trees obtained by the NNI and SPR rearranger. This consensus tree was saved in nexus format and opened in a tree visualizing software FigTree ver. 1.4.3 (Andrew 2016).

Results

The complete cDNA sequence of the *rbcS* gene from diverse organisms belonging to 51 species were selected for the present study. The length of the sequences varied from 342 to 618 nucleotides. Out of 51 sequences, 2 sequences belonged to green non-sulphur bacteria, 5 sequences belonged to Cyanobacteria, 10 sequences belonged to algae, 2 sequences belonged to Bryophytes, 1 sequence each belonged to Pteridophytes & Gymnosperms and 15 sequences each belonged to monocotyledons and dicotyledons (Table 1).

The MSA performed had several gaps/substitutions indicating the magnitude of divergence in the sequences during evolution (Figure. 1). The search for the conserved region in MSA identified one such region between position 218 and 321. The conserved sequence was as follows: 218-TTCGAGACCC TGTCCTACCT GCCACCTCTG ACCGAGGAGC AAATGGCGAA GCAGGTCGAC TACCTGATCC GCAACGGCTG GGTTCCCTGC CTCGAGTTCA AC-321.

Dendrogram drawn using NNI Rearranger

The dendrogram showed two major clades viz., A and B. Out of these two, clade A comprised of 9 taxa from photosynthetic bacteria and algae other than Chlorophyceae. Clade A was further divided into two subclades viz., A1 and A2. The subclade A1 consisted of a single taxon from green non-sulphur bacteria *Kouleothrix*, while the subclade A2 was again divided into two sub-subclades A3 and A4. The sub-subclade A3 comprised of a green non-sulphur bacterium *Nitrolancea*. The sub-subclade A4, on the contrary, consisted of all the algal taxa other than the members of Chlorophyceae. In this sub-subclade, a Dinoflagella – *Karenia*, was out-grouped (A5) while the group A6 consisted of the taxa belonging to Rhodophyceae (*Porphyra*, *Pyropia* and *Saccharina*), Xanthophyceae (*Vaucheria*) and Bacillariophyceae (*Phaeodactylum* and unidentified diatom endosymbiont of *Peridinium foliaceum*) (Figure. 2).

Figure 1: A part of multiple sequence alignment performed using BioEdit

Figure 2: Consensus NNI dendrogram drawn using FigTree

The remaining 42 taxa were included in the clade B. The taxa in this clade segregated in a manner that a single taxon or a group of few taxa separated from the others in each step. The noteworthy feature of this clade was that all the monocotyledons aggregated on a single branch 'O' except *Dioscorea*, *Zantedeschia* and *Wolfia*. These taxa occupied the branch 'P' with all the dicots except *Lotus* and *Vitis*. However, the monocotyledons and dicotyledons clearly segregated over the branches P1 and P2. The remaining dicotyledons viz., *Lotus* and *Vitis* were present in a separate sub-subclade M, where they grouped with the taxa of Chlorophyceae. In the clade B all the taxa of Cyanobacteria were arranged on

Table 1: cDNA sequences of *rbcS* gene retrieved from GenBank for analysis

SN	*Name of the organism*	*Accession number*	*Length of the sequence (nucleotides)*	*Organism*	*Taxonomic group*
1	*Nitrolancea hollandica*	CAGS01000540	424	Prokaryote	Chloroflexi
2	*Kouleothrix aurantiaca*	LJCR01000215	423	Prokaryote	Chloroflexi
3	*Thermoleptolyngbya oregonensis*	EU119379	357	Prokaryote	Cyanobacteria
4	*Anabaena sp.*	U05590	330	Prokaryote	Cyanobacteria
5	*Thermosynechococcus vulcanus*	AB297499	357	Prokaryote	Cyanobacteria
6	*Prochlorothrix hollandica*	X57359	330	Prokaryote	Cyanobacteria
7	*Synechococcus sp.*	U46156	342	Prokaryote	Cyanobacteria
8	*Porphyra yamadae*	LC328319	417	Eukaryote	Rhodophyceae
9	*Pyropia suborbiculata*	AB671542	417	Eukaryote	Rhodophyceae
10	*Saccharina japonica*	HM798587	420	Eukaryote	Rhodophyceae
11	*Vaucheria litorea*	AF207527	420	Eukaryote	Xanthophyceae
12	*Dunaliella salina*	AY646428	573	Eukaryote	Chlorophyceae
13	*Botryococcus braunii*	AB748454	573	Eukaryote	Chlorophyceae
14	*Haematococcus lacustris*	DQ178607	426	Eukaryote	Chlorophyceae
15	*Phaeodactylum tricornutum*	AY819643	420	Eukaryote	Bacillariophyceae
16	*Unidentified diatom endosymbiont of Peridinium foliaceum*	U31876	420	Eukaryote	Bacillariophyceae
17	*Karenia brevis*	KR935866	390	Eukaryote	Dinophyceae
18	*Marchantia polymorpha*	LC064115	546	Eukaryote	Marchantiopsida
19	*Physcomitrella patens*	X76634	231	Eukaryote	Bryopsida
20	*Pteris vittata*	X73136	618	Eukaryote	Polypodiopsida
21	*Larix laricina*	X16039	363	Eukaryote	Pinopsida
22	*Triticum aestivum*	AB042069	525	Eukaryote	Liliopsida
23	*Sorghum bicolor*	AB564718	510	Eukaryote	Liliopsida
24	*Cenchrus purpureus*	LC390054	510	Eukaryote	Liliopsida
25	*Zantedeschia aethiopica*	AF034479	537	Eukaryote	Liliopsida
26	*Dioscorea zingiberensis*	AQM56815	552	Eukaryote	Liliopsida
27	*Phragmites australis*	KF697236	516	Eukaryote	Liliopsida
28	*Oryza sativa*	AY445627	528	Eukaryote	Liliopsida
29	*Echinochloa crus-galli*	KC478601	513	Eukaryote	Liliopsida
30	*Miscanthus x giganteus*	EU219923	510	Eukaryote	Liliopsida
31	*Phleum pratense*	AB976028	528	Eukaryote	Liliopsida
32	*Megathyrsus maximus*	LC390055	510	Eukaryote	Liliopsida
33	*Zea mays*	Y09214	513	Eukaryote	Liliopsida
34	*Aegilops tauschii*	X83095	528	Eukaryote	Liliopsida
35	*Saccharum hybrid cultivar GT28*	JN591757	510	Eukaryote	Liliopsida
36	*Wolffia australiana*	AEJ33935	534	Eukaryote	Liliopsida
37	*Camellia oleifera*	KJ721196	531	Eukaryote	Eudicotyledons
38	*Brassica juncea*	HQ404361	546	Eukaryote	Eudicotyledons

39	*Tragopogon pratensis*	KT879190	540	Eukaryote	Eudicotyledons
40	*Glycine soja*	AAA82069	537	Eukaryote	Eudicotyledons
41	*Lotus japonicus*	LC126623	522	Eukaryote	Eudicotyledons
42	*Nicotiana attenuata*	JF419564	543	Eukaryote	Eudicotyledons
43	*Hevea brasiliensis*	EU449763	552	Eukaryote	Eudicotyledons
44	*Vitis vinifera*	LC126624	510	Eukaryote	Eudicotyledons
45	*Ammopiptanthus mongolicus*	AY843526	525	Eukaryote	Eudicotyledons
46	*Solanum tuberosum*	JX576219	546	Eukaryote	Eudicotyledons
47	*Arachis hypogaea*	AGV08375	549	Eukaryote	Eudicotyledons
48	*Gossypioides kirkii*	JQ034304	505	Eukaryote	Eudicotyledons
49	*Rosa roxburghii*	AEA48975	540	Eukaryote	Eudicotyledons
50	*Eucalyptus globulus*	AB537499	540	Eukaryote	Eudicotyledons
51	*Lactuca sativa*	AF162210	546	Eukaryote	Eudicotyledons

Table 2: Similarities between the trees drawn using NNI and SPR rearranger

1. Separate Chlorophyceae from Rhodophyceae, Xanthophyceae, Bacillariophyceae and Dinoflagellates.
2. Prokaryotes show similarity with Rhodophyceae, Xanthophyceae, Bacillariophyceae and Dinoflagellates.
3. Cyanobacteria more related to Chlorophyceae and plants.
4. Dicots viz., *Lotus* and *Vitis* nearer to Chlorophyceae than angiosperms.
5. All monocots cluster together near dicots.

Table 3: Differences between the trees drawn using NNI and SPR rearranger

NNI Rearranger	*SPR Rearranger*
1. The prokaryotes cluster separately in two groups.	1. The prokaryotes cluster together in a single group.
2. Separates green non-sulphur bacteria from the cyanobacteria.	2. Groups green non-sulphur bacteria and the cyanobacteria.
3. The gymnosperm taxa placed closer to Chlorophyceae, lower plants and two taxa of dicots.	3. The gymnosperm taxa placed between monocots and dicots.
4. Both the bryophytic taxa closely placed.	4. Both the bryophytic taxa distantly placed.

the branches C and E. The branch D included diverse taxa belonging to Bryophytes, Pteridophytes and Gymnosperms (Figure. 2).

Dendrogram drawn using SPR rearranger

The dendrogram showed three major clades 1, 2 and 3. Out of these three, clade 1 comprised of two taxa of Rhodophyceae viz., *Porphyra* and *Pyropia*. The clade 2 consisted of four taxa of algae viz., Rhodophyceae (*Saccharina*), Bacillariophyceae

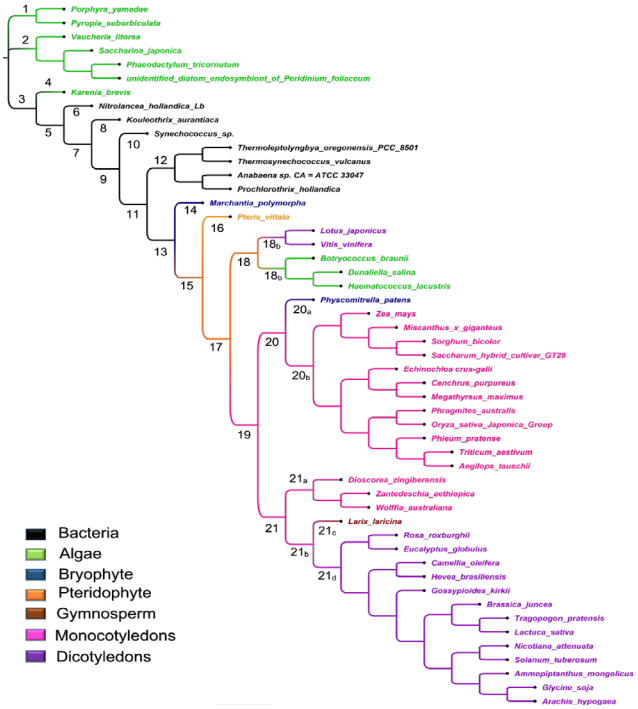

Figure 3: Consensus SPR dendrogram drawn using FigTree

(*Phaeodactylum* and an unidentified diatom endosymbiont of *Peridinium foliaceum*) and Xanthophyceae (*Vaucheria*). The clade 3 was further divided into two subclades 4 and 5. The subclade 4 consisted of a single taxon of Dinoflagellate i.e., *Karenia*. The subclade 5 had the taxa of Chlorophyceae and Cyanobacteria grouped together but clearly segregated over the internal branches. The remaining taxa belonging to Bryophytes, Pteridophytes, Chlorophyceae, Gymnosperms and Angiosperms were grouped together. *Marchantia* (a Bryophyte) and *Pteris* (a Pteridophyte) were placed near to the Cyanobacteria (branches 14 and 16). Strikingly, the Chlorophycean taxa are placed together with the

two taxa of Dicotyledons over branch 18. *Physcomitrella* (a Bryophyte) occupied a sub-subclade 19 consisting of all the Monocotyledons, a Gymnosperm and remaining Dicotyledons (Figure. 3).

Discussion

The present investigation was carried out to study the similarities and differences between the sequence of *rbcS* gene in prokaryotes and eukaryotes. Two types of dendrograms were drawn using NNI rearranger and SPR rearranger. The dendrograms obtained using both these methods revealed a clear-cut distinction between lower and higher photosynthetic organisms. This indicates that *rbcS* gene sequence in bacteria and algae are more alike than the plants, particularly the higher plants. The two dendrograms show similar topology (Table 2) but also reveal some noteworthy distinctions (Table 3). One of the similarities between them is the clustering of three Chlorophycean taxa viz., *Botryococcus, Dunaliella*, and *Haematococcus* in the transition zone over the dendrogram between the clades formed by bacteria and non-green algae and those formed by the plants. This suggests the sequence of *rbcS* gene in Chlorophyceae is similar to plants on one hand, while on the other it resembles with the bacteria and non-green algae. A similar study shows that the *rbcS* gene sequence is analogous in green algae and Cyanobacteria. Similarly, the *rbcS* gene sequence is also analogous in non-green algae and other bacteria. Likewise, there is a divergence in the gene sequence of green and non-green algae (Assali 1991).

Further both the rearrangers have kept bacteria closer to the plants than the non-green algae. This indicates that the sequence of *rbcS* gene in bacteria is considerably different than the non-green algae. The phylogenetic studies of *rbcS* gene in bacteria, algae and plants reveal that the Chlorophyceaen taxa and Cyanobacteria are more related to plants. In addition, the Gymnosperms displays distinct *rbcS* gene sequence as compared to Angiosperms (Pasternak and Glick 1992). In the present study also the lone gymnosperm *Larix laricina* clustered differently on the NNI and SPR tree. This probably might be due to its unique nucleotide sequence which was in a larger part different than the other sequences.

Conclusion

Thus, from the findings of the present investigation it can be concluded that there is a marked difference in the nucleotide sequence of *rbcS* gene between the higher plants and the lower plants along with the algae and photosynthetic prokaryotes. Moreover, the findings of a tool used to construct the dendrogram should be validated with other methods to draw a reliable inference from the molecular data.

References:

Andrew R (2016). FigTree: tree figure drawing tool, version 1.4.3. Institute of Evolutionary Biology University of Edinburgh, United Kingdom.

Assali NE, Martin WF, Sommerville CC and Loiseaux-de Goër S (1991). Evolution of the Rubisco operon from prokaryotes to algae: structure and analysis of the *rbcS* gene of the brown alga *Pylaiella littoralis*. *Plant Molecular Biology* **17(4):** 853-863.

Erb TJ and Zarzycki J (2018). A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme. *Curr. Opinion in Biotechnology* **49:** 100-107.

Goldschmidt-Clermont M and Rahire M (1986). Sequence, evolution and differential expression of the two genes encoding variant small subunits of ribulose bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *Journal of Molecular Biology.* 191(3): 421-432.

Hall T (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic Acids Symp. Ser. **41:** 95-98.

Jarvis P and Soll J (2001). Toc, Tic, and chloroplast protein import. *Biochimica et Biophysica Acta.* **1541(1-2):** 64-79.

Pasternak J J and Glick B R (1992). Molecular evolutionary analyses of the small and large subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Canadian J. Bot.* **70(4):** 715-723.

Razzak M, Lee D W, Yoo Y J and Hwang I (2017). Evolution of rubisco complex small subunit transit peptides from algae to plants. *Scientific reports* **7(1):** 1-11.

Suzuki Y, Nakabayashi K, Yoshizawa R, Mae T and Makino A (2009). Differences in expression of the RBCS multigene family and Rubisco protein content in various rice plant tissues at different growth stages. *Plant and Cell Physiology* **50(10):** 1851-1855.

Xiang F, Fang Y and Xiang J (2016). Structural and evolutionary relationships among RuBisCOs inferred from their large and small subunits. *Zeitschrift für Naturforschung C.* **71(5-6):** 181-189.