www.rjlbpcs.com



Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/



Original Research Article

DOI: 10.26479/2019.0502.16

PHYLOGENETIC UTILITY OF GENE ORDER AND SYNTENY AMONG THE CHLOROPLAST GENOMES OF MICROALGAE

Suneetha Saragadam^{1*}, Uma Devi Koduru¹, Prajna Muthabathula^{1,2}

1. Department of Botany, Andhra University, Visakhapatnam, Andhra Pradesh, India.

2. Department of Botany, AdikaviNannaya University, Rajamahendravaram, Andhra Pradesh, India.

ABSTRACT: Chloroplast plays a major role in photosynthesis due to small size genome and it houses a small number of genes compared to the nuclear genome. The present study compares the gene order and synteny among the microalgae. A phylogenetic tree is inferred from gene order data. The synteny analysis specifies that microalgae undergo many genome rearrangements and high variability in gene order between and within phyla. Phylogenetic tree from gene order data is applicable to class level.

KEYWORDS: chloroplast genome, gene order, conserved, synteny, phylogeny.

Corresponding Author: Suneetha Saragadam*

Department of Botany, Andhra University, Visakhapatnam, Andhra Pradesh, India Email Address: chand.sunitha@gmail.com

1. INTRODUCTION

Genome data is accumulating with the available genome sequencing strategies which are aiding the comparative genomics. Phylogenetic relationships are studied based on two types of data i.e., gene order data and sequence data. The first kind of data relates to the order of genes on the chromosomes whereas the second kind relates to the gene (DNA) sequence. All these categories were analysed for their conservatism at different levels in the genome of related species or organisms. Synteny blocks is another kind which contains the genes that share similar functions and are mostly co-regulated [1], which are much more complex in eukaryotes than the prokaryotes. Poor conservation of gene order in prokaryotes provides better for analysis of the species at strain level evolution [2]. As the prokaryotic genes are regulated by a single regulatory element called operons which are absent in the higher plant and animal genomes with some exceptions. The accumulating sequenced data

Saragadam et al RJLBPCS 2019 www.rjlbpcs.com Life Science Informatics Publications suggests that due to a functional role, the genes are arranged in a cluster called gene clusters in eukaryotes [3]. Gene order data is based on structural arrangements of genes in an organism's entire genome. Phylogenetic reconstruction based on gene order data which determines the evolution of the organism. Evolutionary process alters the order of genes through rearrangements like inversions and transpositions, collectively known as genome rearrangements. Such events are rare so, phylogenies were inferred using rearrangement of gene orders [4-10]. Synteny blocks help in knowing the gene functions and they tend to be conserved in related organisms. However, differences in the gene order may be observed in genomes. Phylogenetic inference based on gene order data resolve the species relationship in mitochondria and chloroplast genome [6, 7, 11]. As chloroplast contains small genome (single chromosome) it is evolving at a faster rate as a result order of genes undergo insertions, inversions and deletions, also transposition (in mitochondria) and translocation between chromosomes during evolution. Such differences in gene order provide insights in inferring their phylogenetic relationships among microalgae. Gene order data from plastid genomes proven useful in resolving phylogenetic relationships among plant groups [12]. When gene order data from close relatives is taken the least number of changes was observed [13]. Highly rearranged chloroplast genome is noticed in Chalymadomonas reinhardtii when compared with other green algae [14]. In our analysis, preparation of gene order data from chloroplast genome of microalgae which is used to infer phylogeny to resolve the relationship among microalgae and synteny analysis is done to find the synteny blocks shared in microalgae.

2. MATERIALS AND METHODS

1. Chloroplast genome sequence retrieval

Chloroplast genome files available to date were retrieved from NCBI ftp site (ftp;//ftp.ncbi.nlm.nih.gov/genomes/archive/old refseq) for all 43 microalgal species. We analysed thirty protein coding genes that are shared by at least 43 taxa.

2. Gene order analysis

Pairwise comparison of gene order depends upon the type of the comparator chosen. No synteny is observed when the distant comparator is chosen. In our case,*Monomorphina aenigmatica* (Euglenophyta) is chosen as a comparator as it consists of less number of genes. In preparation of data set where n denotes any number ranging from 1-30 and –n or n denotes the minus or plus strand where the gene is present. Genes are ordered based on their location irrespective of their presence in plus or minus strands [15].

3. Phylogenetic analysis

Phylogenetic constructions from gene order data were performed using MLGO (Maximum Likelihood for gene order analysis) a web server http://www.geneorder.org/ [16].

4. Synteny analysis

Synteny analysis was performed using Progressive mauve software © 2019 Life Science Informatics Publication All rights reserved Peer review under responsibility of Life Science Informatics Publications 2019 March – April RJLBPCS 5(2) Page No.202

3. RESULTS AND DISCUSSION

1. Gene content and genome organization

Chloroplast genomes of microalgae are 60-269 kb in length (Table 1). Smallest plastid genome of microalgae is reported in *Eutreptiella gymnastica*. The largest genome is from a Chlorophycean member - *Dunaliella* [18]. Gene number in the genomes range from 92-243, these include rRNA, 26-37 tRNA, 61-207 protein coding genes.GC% range from 26.1 (*Euglena gracilis*) to 57.7 (*Trebouxiophyceae*) (Table 1). Presence or absence of introns varies from one group to next.

S. No	RefSeq	Organism	Phylum	Gen	ome	No.	of			
						Intr	ons			
						Gro	up			
				Size		Ι	II			
				(MB)	GC%					
1	NC_014267	Kryptoperidinium	Dinophyta	0.14	32.4	0	0			
		foliaceum	(Dinophyceae)							
2	NC_014287	Durinskia baltica	Dinophyta	0.11	32.6	0	0			
			(Dinophyceae)							
3	NC_001713	Odontella sinensis	Heterokontophyta	0.12	31.8	0	0			
			(Bacillariophycea)		2					
4	NC_008588	Phaeodactylum	Heterokontophyta	0.12	32.5	0	0			
		tricornutum	(Bacillariophycea)		6					
5	NC_008589	Thalassiosira	Heterokontophyta	0.13	30.6	0	0			
		pseudonana	(Bacillariophycea)		6					
6	NC_014808	Thalassiosira	Heterokontophyta	0.12	34	1	0			
		oceanica	(Bacillariophycea)							
7	NC_015403	Fistulifera	Heterokontophyta	0.13	32.1	0	0			
			(Bacillariophycea)		3					
8	NC_010772	Heterosigma	Heterokontophyta	0.16	30.4	0	0			
		akashiwo	(Raphidophyceae)		4					
9	NC_012898	Aureococcus	Heterokontophyta	0.09	33.6	0	0			
		anophagefferens	(Pelagophyceae)		4					
10	NC_012903	Aureoumbra	Heterokontophyta	30	0	0				
		lagunensis	(Pelagophyceae)							

Table 1. General features of microalgae

Saraga	dam et al RJLBP	CS 2019 www	rjlbpcs.com L	ife Science	Informa	tics Pu	ublication		
11	NC_020014	Nannochloropsis	Heterokontophyta	0.11	32.9	0	0		
		gaditana	(Eustigmatophyceae)		5				
12	NC_016703	Phaeocystis	Haptophyta	0.11	35.5	0	0		
		antarctica	(Coccolithophyceae)		1				
13	NC_021637	Phaeocystis globosa	Haptophyta	0.11	35.4	0	0		
			(Coccolithophyceae)		2				
14	NC_007288	Emiliania huxleyi	Haptophyta	0.1	36.8	0	0		
			(Prymnesiophyceae)						
15	NC_020371	Pavlova lutheri	Haptophyta	0.1	35.5	0	0		
			(Pavlovophyceae)		9				
16	NC_000926	Guillardia theta	Cryptophyta	0.12	32.9	0	0		
			(Cryptophyceae)		7				
17	NC_009573	Rhodomonas salina	Cryptophyta	0.14	34.7	0	2		
			(Cryptophyceae)		9				
18	NC_001840	Cyanidium caldarium	Rhodophyta	0.16	32.7	0	0		
			(Cyanidiophyceae)		3				
19	NC_004799	Cyanidioschyzon	Rhodophyta	0.15	37.6	0	0		
		merolae strain 10D	(Cyanidiophyceae)		3				
20	NC_001675	Cyanophora	Glaucophyta	0.13	30.5	3	0		
		paradoxa	(Glaucophyceae)						
21	NC_001603	Euglena gracilis	Euglenophyta	0.14	26.1	160	0		
			(Euglenophyceae)						
22	NC_020460	Euglena viridis	Euglenophyta	0.09	26.4	77	0		
			(Euglenophyceae)						
23	NC_020018	Monomorphina	Euglenophyta	0.07	29.4	53	0		
		aenigmatica	(Euglenophyceae)						
24	NC_017754	Eutreptiella	Euglenophyta	0.06	34.3	2	0		
		gymnastica	(Euglenophyceae)						
25	NC_012099	Pyramimonas	Chlorophyta	0.1	34.7	0	1		
		parkeae	(Pyramimonadophyceae))					
26	NC_008289	Ostreococcus tauri	Chlorophyta	0.07	39.8	0	1		
			(Prasinophyceae)		9				
27	NC_012097	Pycnococcus	Chlorophyta	0.08	39.5	0	1		
		provasolii	(Pyramimonadophyceae))					
	1								

Saragao	dam et al RJLBP	CS 2019 www	rjlbpcs.com Life	Science	Informa	tics Pı	ublication		
28	NC_000927	Nephroselmis	Chlorophyta	0.2	42.1	0	0		
		olivacea	(Nephrophyceae)		4				
29	NC_012101	Monomastix sp.	Chlorophyta	0.11	39	5	1		
		OKE-1	(Prasinophyceae)						
30	NC_005353	Chlamydomonas	Chlorophyta	0.2	34.4	5	2		
		reinhardtii	(Chlorophyceae)		6				
31	NC_008101	Scenedesmus	Chlorophyta	0.16	26.8	7	2		
		obliquus	(Chlorophyceae)		9				
32	NC_016732	Dunaliella salina	Chlorophyta	0.27	32.0	35	8		
			(Chlorophyceae)		8				
33	NC_016733	Pedinomonas minor	Chlorophyta(Ulvophyceae)	0.09	34.8	0	0		
34	NC_008099	Oltmannsiellopsis	Chlorophyta	0.15	40.4	5	0		
		viridis	(Ulvophyceae)		7				
35	NC_008114	Pseudendoclonium	Chlorophyta	0.2	31.4	27	0		
		akinetum	(Ulvophyceae)		9				
36	NC_001865	Chlorella vulgaris	Chlorophyta	0.15	31.6	3	0		
			(Trebouxiohyceae)						
37	NC_015359	Chlorella variabilis	Chlorophyta	0.12	34	3	0		
			(Trebouxiohyceae)						
38	NC_012978	Parachlorella	Chlorophyta	0.12	30	1	0		
		kessleri	(Trebouxiohyceae)						
39	NC_015084	Соссотуха	Chlorophyta	0.17	50.7	0	1		
		subellipsoidea C-169	(Trebouxiohyceae)						
40	NC_018569	Trebouxiophyceae sp.	Chlorophyta	0.14	57.7	5	0		
			(Trebouxiohyceae)						
41	NC_002186	Mesostigma viride	Streptophyta	0.11	30	0	0		
			(Mesostigmatophyceae)						
42	NC_004115	Chaetosphaeridium	Streptophyta	0.13	29.6	1	17		
		globosum	(Coleochaetophyceae)		2				
43	NC_008116	Staurastrum	Streptophyta	0.16	32.4	1	7		
		punctulatum	(Zygnemophyceae)		9				

*classes of the microalgae are shown in brackets

Introns are termed as the non-coding region between the exons within a single gene. Out of four types of introns, but only two types (group I and group II) occur in organelle genomes [19]. Introns are completely absent in plastid genomes in members of Dinophyta, Haptophyta, Rhodophyta and

Saragadam et al RJLBPCS 2019www.rjlbpcs.comLife Science Informatics PublicationsHeterokontophyta (except in Thalassiosira oceanica). Euglena is overloaded with atleast 149 introns[20]. Prasinophyta clade is composed of only group II introns. UTC (Ulvophyceae,Trebouxiophyceae, Chlorophyceae) clade is majorly interrupted with the group I introns (Table 1).

2. Gene order of core protein coding genes in the chloroplast genome

Majority of the orthologous protein coding genes are present in the plus strand of the chloroplast genome (Table 2). Some orthologous gene clusters are interrupted by genes in some microalgae.

Damba balania 9 9 9 9 9 9 9 9 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0<	Kryptoperidinium foliaceum	-10	-9	5	6	-8	12	-11	13	-14	-3	-2	-23	30	16	-15	4	1	28	25	24	17	27	26	20	29	21	22	19	18	-7
Decombe Decombe <t< th=""><th>Durinskia baltica</th><th>6</th><th>5</th><th>9</th><th>10</th><th>-8</th><th>12</th><th>-11</th><th>13</th><th>-14</th><th>-3</th><th>-2</th><th>-23</th><th>30</th><th>16</th><th>-15</th><th>4</th><th>1</th><th>-28</th><th>7</th><th>-18</th><th>25</th><th>24</th><th>17</th><th>27</th><th>26</th><th>20</th><th>29</th><th>21</th><th>22</th><th>19</th></t<>	Durinskia baltica	6	5	9	10	-8	12	-11	13	-14	-3	-2	-23	30	16	-15	4	1	-28	7	-18	25	24	17	27	26	20	29	21	22	19
Pace-octa 6 5 5 6 1 0 0 0 0 <th>Odontella sinensis</th> <th>-19</th> <th>-22</th> <th>-21</th> <th>-29</th> <th>-20</th> <th>-26</th> <th>-27</th> <th>-17</th> <th>-24</th> <th>-25</th> <th>18</th> <th>-7</th> <th>-28</th> <th>-15</th> <th>4</th> <th>1</th> <th>-8</th> <th>14</th> <th>-13</th> <th>11</th> <th>12</th> <th>9</th> <th>10</th> <th>-3</th> <th>-2</th> <th>-23</th> <th>30</th> <th>16</th> <th>-6</th> <th>-5</th>	Odontella sinensis	-19	-22	-21	-29	-20	-26	-27	-17	-24	-25	18	-7	-28	-15	4	1	-8	14	-13	11	12	9	10	-3	-2	-23	30	16	-6	-5
Tablesionis secondes 6 5 6 7 8 7 8 5 4 7 8 5 4 7 8 5 4 7 8 5 7 8 5 7 8 5 7 8 5 7 8 5 7 8 5 7 8 5 7 8 5 7 8 7 8 7 8 7 8 7 8 5 8 7 8 5 8 7 8 5 8 7 8 5 8 7 8 5 8 7 8	Phaeodactylum tricornutum	6	-5	9	10	-8	12	-11	13	-14	-3	-2	-23	30	16	-15	4	1	28	7	-18	25	24	17	27	26	20	29	21	22	19
Thansshinger 16 00 0 0 0 0 </th <th>Thalassiosira pseudonana</th> <th>6</th> <th>5</th> <th>-15</th> <th>4</th> <th>1</th> <th>14</th> <th>-13</th> <th>11</th> <th>28</th> <th>-16</th> <th>-30</th> <th>23</th> <th>2</th> <th>3</th> <th>-12</th> <th>8</th> <th>-10</th> <th>-9</th> <th>7</th> <th>-18</th> <th>25</th> <th>24</th> <th>17</th> <th>27</th> <th>26</th> <th>20</th> <th>29</th> <th>21</th> <th>22</th> <th>19</th>	Thalassiosira pseudonana	6	5	-15	4	1	14	-13	11	28	-16	-30	23	2	3	-12	8	-10	-9	7	-18	25	24	17	27	26	20	29	21	22	19
Findem 20 2 </th <th>Thalassiosira oceanica</th> <th>-16</th> <th>-30</th> <th>23</th> <th>2</th> <th>3</th> <th>9</th> <th>10</th> <th>8</th> <th>12</th> <th>14</th> <th>-13</th> <th>11</th> <th>28</th> <th>-15</th> <th>4</th> <th>1</th> <th>-6</th> <th>5</th> <th>-7</th> <th>18</th> <th>25</th> <th>24</th> <th>17</th> <th>27</th> <th>26</th> <th>20</th> <th>29</th> <th>21</th> <th>22</th> <th>19</th>	Thalassiosira oceanica	-16	-30	23	2	3	9	10	8	12	14	-13	11	28	-15	4	1	-6	5	-7	18	25	24	17	27	26	20	29	21	22	19
Interestigna adaption 5 60 7	Fistulifera	23	2	3	28	-19	-22	-21	-29	-20	-26	-27	-17	-24	-25	18	-7	-6	-5	9	10	-8	12	-11	13	-14	15	4	1	16	30
Aute Constant Constant <thconstant< th=""> Constant <thc< th=""><th>Heterosigma akashivo</th><th>15</th><th>-8</th><th>9</th><th>10</th><th>12</th><th>18</th><th>5</th><th>6</th><th>28</th><th>-7</th><th>-16</th><th>-30</th><th>23</th><th>4</th><th>1</th><th>-3</th><th>-2</th><th>19</th><th>-22</th><th>-21</th><th>-29</th><th>-20</th><th>-26</th><th>-27</th><th>-17</th><th>-24</th><th>-25</th><th>11</th><th>13</th><th>-14</th></thc<></thconstant<>	Heterosigma akashi v o	15	-8	9	10	12	18	5	6	28	-7	-16	-30	23	4	1	-3	-2	19	-22	-21	-29	-20	-26	-27	-17	-24	-25	11	13	-14
Auroenduropic gadiana 5 6 7	Aureococcus anophagefferens	12	28	30	16	4	1	-8	25	24	17	27	26	20	29	21	22	19	14	-13	11	-10	-9	5	6	15	18	23	2	3	7
Namechologis gadiana 5 6 -2 -1 1 1 1 2	Aureoumbra lagunensis	8	25	24	27	17	26	20	29	21	22	19	14	-13	11	-10	-9	5	6	15	30	-18	-23	16	4	1	-3	-2	-12	28	7
Phaeeogusis antaretica 23 37 25 24 47 27 25 24 27 25 24 27 25 24 27 25 24 27 25 24 27 25 26 25 6 6 6 6 6 6 1 4 1<	Nannochloropsis gaditana	5	6	-12	-11	13	-14	-19	-22	-21	-29	-20	-26	-27	-17	-24	-25	8	-28	9	10	-18	7	-15	-3	-23	30	16	4	1	-2
Phacecysis globosa 23 7 25 24 7 27 25 20 23	Phaeocystis antarctica	23	-7	25	24	17	27	26	20	29	21	22	19	5	6	-3	-2	-12	14	-13	11	9	10	8	-1	-4	-16	-18	-30	15	-28
Entania bandeyi 5 6 7	Phaeocystis globosa	23	-7	25	24	17	27	26	20	29	21	22	19	5	6	-3	-2	-12	14	-13	11	9	10	8	-1	-4	-16	-18	-30	15	-28
Pathons batheri C C <	Emiliania huxleyi	5	6	15	7	-28	23	-3	-2	-12	14	-13	11	30	18	16	4	1	-8	-10	-9	25	24	17	27	26	20	29	21	22	19
Guilardia theta -6 -7 -8 -7 -8 -9 -8	Pavlova lutheri	25	24	17	27	26	20	15	14	-13	11	18	12	-19	-22	-21	-29	-8	-16	-30	28	-10	-9	-1	-4	23	7	5	6	-3	-2
Phodomonas saina 96 97 98 97 98 97	Guillardia theta	-8	-7	18	15	-10	-9	7	-4	-16	-30	23	2	3	-28	12	-6	-5	25	24	17	27	26	20	29	21	22	19	14	-13	11
Cpanidium caldarium 6 2 3 -23 3 15 4 1 10 5 7 11 1 10 10 10 10 2 2 2 2 2 2 2 10 3 2 4 1 10 3 10 5 6 2 2 10 2 2 10 2 2 10 2 2 3 4 1 10 3 10 2 <th2< th=""> 2 <th2< th=""></th2<></th2<>	Rhodomonas salina	-8	-7	18	15	-10	-9	-1	-4	-16	-30	23	2	3	-28	12	-6	-5	25	24	17	27	26	20	29	21	22	19	14	-13	11
Cyanolicochycon meroles strain 10 5 7 8 14 70 70 70 <th< th=""><th>Cyanidium caldarium</th><th>8</th><th>2</th><th>3</th><th>-23</th><th>30</th><th>16</th><th>4</th><th>1</th><th>-10</th><th>-9</th><th>7</th><th>-11</th><th>13</th><th>14</th><th>18</th><th>15</th><th>-19</th><th>-22</th><th>-21</th><th>-29</th><th>-20</th><th>-26</th><th>-27</th><th>-17</th><th>-24</th><th>-25</th><th>-28</th><th>12</th><th>-6</th><th>-5</th></th<>	Cyanidium caldarium	8	2	3	-23	30	16	4	1	-10	-9	7	-11	13	14	18	15	-19	-22	-21	-29	-20	-26	-27	-17	-24	-25	-28	12	-6	-5
Canophora paradoxa 5 6 2 1 1 1 2 2 2 1 2 2 1 4 4 6 3 0 5 2	Cyanidioschyzon merolae strain 10D	18	15	7	8	14	-13	11	5	6	-12	28	25	24	17	27	26	20	29	21	22	19	-3	-2	-24	30	16	4	1	-10	-9
Euglena graciiis 8 7 -12 -13 -22 -23 5 6 9 10 25 24 17 27 28 20 28 6 4 1 23 -55 3 21 30 10 25 24 17 27 28 20 28 6 4 1 23 -55 3 2 30 30 3 55 25 3 20 28 20 28 20 25 25 3 3 3 3 5 25 3 5	Cyanophora paradoxa	5	6	28	13	14	11	9	10	-20	-26	-27	17	-24	-25	23	19	-22	18	-1	-4	-16	-30	-15	7	2	3	8	-12	-21	-29
Euglena windis 8 7 -12 9 9 1 2 2 3 1 1 2 2 3 1 1 1 2 1 1 1 1 2 1 <th1< th=""> 1 1</th1<>	Euglena gracilis	8	7	-12	-19	-22	-28	5	6	9	10	25	24	17	27	26	20	29	16	4	1	23	-15	-3	-2	13	11	-14	18	21	-30
Monomorphina aenigmatica 8 7 -12 2 1 -13 2 3 -15 -23 -1 -4 -16 -23 -20 -26 -20 -20 -20 <	Euglena viridis	8	7	-12	-19	-22	-28	5	6	3	10	25	24	17	27	26	20	29	16	4	1	23	-15	-3	-2	13	11	-14	18	21	-30
Eutropiella gymnastica 2 3 5 6 7 <th>Monomorphina aenigmatica</th> <th>8</th> <th>7</th> <th>-12</th> <th>21</th> <th>-30</th> <th>14</th> <th>-11</th> <th>-13</th> <th>2</th> <th>3</th> <th>-15</th> <th>-23</th> <th>-1</th> <th>-4</th> <th>-16</th> <th>-29</th> <th>-20</th> <th>-26</th> <th>-27</th> <th>-17</th> <th>-24</th> <th>-25</th> <th>-10</th> <th>-9</th> <th>-6</th> <th>-5</th> <th>18</th> <th>28</th> <th>22</th> <th>19</th>	Monomorphina aenigmatica	8	7	-12	21	-30	14	-11	-13	2	3	-15	-23	-1	-4	-16	-29	-20	-26	-27	-17	-24	-25	-10	-9	-6	-5	18	28	22	19
Pyraminons parkee 7 11 13 14 10 9 12 23 20	Eutreptiella gymnastica	2	3	15	25	24	17	27	26	20	29	7	-12	8	28	25	19	5	6	9	10	23	-21	-18	-30	13	11	-14	16	4	1
Ostrecoccus tauri -8 -4 -23 -3 -2 -10 -3 -10 23 22 19 30 16 25 24 17 27 26 20 21 4 1 -6 55 28 12 15 7 Pycnocccus provasolii -15 -30 2 3 90 -22 25 1 -4 -20 26 21 18 2 10 -8 -23 -8 -23 -11 13 -14 5 6 -8 -23 -10 -8 -23 -10 -8 -23 -10 -8 -23 -10 -13 -14 -3 -2 -16 -8 -21 -11 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 -14 -13 <th>Pyramimonas parkeae</th> <th>7</th> <th>-11</th> <th>13</th> <th>-14</th> <th>-10</th> <th>-9</th> <th>-12</th> <th>23</th> <th>-21</th> <th>-18</th> <th>-29</th> <th>-20</th> <th>-26</th> <th>-27</th> <th>-17</th> <th>-24</th> <th>-25</th> <th>-15</th> <th>2</th> <th>3</th> <th>-6</th> <th>5</th> <th>16</th> <th>-28</th> <th>22</th> <th>19</th> <th>-1</th> <th>-4</th> <th>-30</th> <th>16</th>	Pyramimonas parkeae	7	-11	13	-14	-10	-9	-12	23	-21	-18	-29	-20	-26	-27	-17	-24	-25	-15	2	3	-6	5	16	-28	22	19	-1	-4	-30	16
Pycnococcus provasoli -15 -30 2 3 9 10 -12 22 19 4 1 25 24 17 27 28 20 23 21 8 -11 13 -14 7 -18 5 6 -28 23 18 Nephroselmis olivacea 7 22 3 -21 -22 5 6 -28 -27 -77 -24 -25 -1 -4 30 16 8 20 -17 -12 14 13 -14 23 24 77 -24 -25 14 -13 -14 -10 -14 -14 -13 -14 -14 -14 -13 -14 -13 -14 <th>Ostreococcus tauri</th> <th>-8</th> <th>-14</th> <th>-29</th> <th>-3</th> <th>-2</th> <th>-10</th> <th>-9</th> <th>-11</th> <th>13</th> <th>-18</th> <th>23</th> <th>22</th> <th>19</th> <th>30</th> <th>16</th> <th>25</th> <th>24</th> <th>17</th> <th>27</th> <th>26</th> <th>20</th> <th>21</th> <th>4</th> <th>1</th> <th>-6</th> <th>-5</th> <th>28</th> <th>12</th> <th>15</th> <th>7</th>	Ostreococcus tauri	-8	-14	-29	-3	-2	-10	-9	-11	13	-18	23	22	19	30	16	25	24	17	27	26	20	21	4	1	-6	-5	28	12	15	7
Nephroselmis olivacea 7 2 3 -21 -20	Pycnococcus provasolii	-15	-30	2	3	9	10	-12	22	19	4	1	25	24	17	27	26	20	29	21	8	-11	13	-14	7	-18	5	6	-28	23	16
Monomastik sp. OKE-1 8 -19 -22 5 6 -28 75 -30 -20 -1 -4 -3 -2 -16 93 10 -18 -11 -12 14 13 25 24 17 27 26 20 23 -7 Chanydomonas reinhardii 12 -28 23 25 24 27 26 20 5 18 -13 14 16 -28 49 10 17 30 6 -5 1 4 21 7 24 25 18 23 27 26 23 7 24 25 18 23 27 26 23 7 12 24 13 14 13 14 23 17 24 25 18 23 25 18 23 27 26 23 27 26 23 27 25 23 23 25 24 17 27 26 23 27 23 23 23 23 23 23	Nephroselmis olivacea	7	2	3	-21	-29	-20	-26	-27	-17	-24	-25	-1	-4	30	16	8	22	19	12	-10	-9	-23	-18	-28	-11	13	-14	5	6	-15
Chanydomonas reinhardtii 12 -28 29 25 24 27 26 20 5 18 -3 -19 -11 13 -14 16 -28 49 10 17 30 6 -15 1 4 21 7 -2 -22 8 Scenedesmus obliquus 5 -4 -30 17 10 -22 -8 -3 -19 16 14 13 -11 12 -21 -15 28 -7 -1 -24 -25 18 -23 -2 -8 -20 -26 -27 -10 -24 -25 18 -23 -2 -8 -23 -7 -10 13 -14 -16 -23 -7 -10 -16 -23 -2 -23 -23 -23 -24 -25 18 -3 -16 -13 3 -15 11 41 16 -23 -2 -23 -23 -23 -23 -23 -23 -23 -23 -23 -23 -23 -24 </th <th>Monomastix sp. OKE-1</th> <th>8</th> <th>-19</th> <th>-22</th> <th>5</th> <th>8</th> <th>-28</th> <th>15</th> <th>-30</th> <th>-29</th> <th>-1</th> <th>-4</th> <th>-3</th> <th>-2</th> <th>-16</th> <th>9</th> <th>10</th> <th>-18</th> <th>-21</th> <th>-11</th> <th>-12</th> <th>14</th> <th>13</th> <th>25</th> <th>24</th> <th>17</th> <th>27</th> <th>26</th> <th>20</th> <th>23</th> <th>-7</th>	Monomastix sp. OKE-1	8	-19	-22	5	8	-28	15	-30	-29	-1	-4	-3	-2	-16	9	10	-18	-21	-11	-12	14	13	25	24	17	27	26	20	23	-7
Scenedesmus obliques 5 -4 -30 17 10 -22 -8 -3 -19 16 14 12 11 12 21 -5 28 -7 -1 -24 -25 18 -23 -2 -8 -20 -26 -27 Dunalela saina 27 28 20 5 -23 -22 -28 -18 28 -17 -24 -25 18 -3 -18 -	Chlamydomonas reinhardtii	12	-28	29	25	24	27	26	20	5	18	-3	-19	-11	13	-14	-16	-28	-9	10	17	30	6	-15	1	4	21	7	-2	-22	8
Dunalella salina 27 26 20 5 -22 -8 -18 28 12 1 25 24 2 4 21 8 -18 3 -15 49 -11 13 -14 -16 -23 -7 10 17 30 Pedinomonas minor 1 -4 -16 -30 14 -13 11 -21 -23 -20 -26 -5 10 -5 12 -9 -10 -15 2 3 18 8 -7 Oltmannsiellopsis viridis -11 13 14 25 24 17 27 26 20 29 21 15 20 18 -3 -2 -6 -5 10 -9 -10 12 3 18 8 -7 Pseudendoclonium akinetum 30 25 24 17 27 26 20 29 21 15 -66 -72 30 15 -83 2 -21 -4 -16 -65 -65 10 <t< th=""><th>Scenedesmus obliguus</th><th>5</th><th>-4</th><th>-30</th><th>17</th><th>10</th><th>-22</th><th>-8</th><th>-3</th><th>-19</th><th>16</th><th>14</th><th>13</th><th>11</th><th>12</th><th>-21</th><th>-15</th><th>28</th><th>-7</th><th>-1</th><th>-24</th><th>-25</th><th>18</th><th>-23</th><th>-2</th><th>-6</th><th>-29</th><th>-9</th><th>-20</th><th>-26</th><th>-27</th></t<>	Scenedesmus obliguus	5	-4	-30	17	10	-22	-8	-3	-19	16	14	13	11	12	-21	-15	28	-7	-1	-24	-25	18	-23	-2	-6	-29	-9	-20	-26	-27
Pedinomonas minor 1 -4 -16 -30 14 -13 11 -21 -23 -20	Dunaliella salina	27	26	20	5	-29	-22	-8	-18	28	12	1	25	24	2	4	21	6	-19	3	-15	-9	-11	13	-14	-16	-23	-7	10	17	30
Oltmannsiellopsis windis -11 13 14 25 24 17 27 26 20 23 21 -12 15 -30 18 -33 -2 -6 -5 -10 -83 -10 -43 -16 -23 -28 8 7 Pseudendoclonium akinetum 30 25 24 17 27 26 20 29 21 15 2 3 7 -19 -22 8 4 13 11 9 10 12 -5 -1 -4 -16 -6 -6 -23 28 14 13 11 9 10 12 -5 -1 -4 -16 -6 -6 -7	Pedinomonas minor	1	-4	-16	-30	14	-13	11	-21	-29	-20	-26	-27	-17	-24	-25	-19	-22	-23	-28	-6	-5	12	-9	-10	-15	2	3	18	8	-7
Pseudendoclonium akinetum 30 25 24 17 27 26 20 28 21 15 2 3 7 -13 -22 8 -13 -23 -23 14 13 11 9 100 12 -5 -1 -4 -16 -8 Chlorella vulgaris -7 -1 -4 -16 -8 28 23 22 19 -10 -5 6 -12 30 15 -8 3 2 -21 -23 20 -23 20 -28 21 27 26 20 23 21 44 16 4 1 24 -5 66 -12 30 15 43 2 -21 23 20 28 21 23 22 13 23 23 24 17 27 26 20 23 21 84 13 11 20 23 22 13 13 14 13 11 20 23 22 13 13 14 13	Oltmannsiellopsis viridis	-11	13	14	25	24	17	27	26	20	29	21	-12	15	-30	18	-3	-2	-6	-5	-10	-9	-19	-22	-1	-4	-16	-23	-28	8	7
Chlorella vulgaris -7 -1 -4 -16 -8 28 22 19 -10 -5 -11 13 -14 5 66 -72 30 15 -18 -32 2 -70 -72 -70 -28 -20 -28 -20 -28 -27 -77 -74 -25 Chlorella variabilis 3 10 -12 5 66 7 -15 -30 14 -13 11 28 23 22 19 25 24 17 27 26 20 23 21 8 4 16 4 1 2 3 18 Parachlorella kessleri -18 -3 -2 -5 18 -10 -30 14 -13 11 20 23 22 19 25 24 17 27 26 20 23 21 8 4 16 1 15 7 Cocomyas subellipsoidea C-165 4 1 12 7 74 1 12 8 <t< th=""><th>Pseudendoclonium akinetum</th><th>30</th><th>25</th><th>24</th><th>17</th><th>27</th><th>26</th><th>20</th><th>29</th><th>21</th><th>15</th><th>2</th><th>3</th><th>7</th><th>-19</th><th>-22</th><th>8</th><th>-18</th><th>-23</th><th>-28</th><th>14</th><th>13</th><th>11</th><th>9</th><th>10</th><th>12</th><th>-5</th><th>-1</th><th>-4</th><th>-16</th><th>-6</th></t<>	Pseudendoclonium akinetum	30	25	24	17	27	26	20	29	21	15	2	3	7	-19	-22	8	-18	-23	-28	14	13	11	9	10	12	-5	-1	-4	-16	-6
Chlorella variabilis 3 10 -12 5 6 7 -15 -33 14 -13 11 23 24 17 27 26 20 23 21 8 16 4 1 2 3 18 Parachlorella kessleri -18 -3 -2 -6 -5 12 9 -30 14 -13 11 28 23 22 19 25 24 17 27 26 20 23 21 8 16 1 15 7 Coccomyna subellipsoidea C-169 4 1 12 7 -8 -3 -2 5 8 -16 21 -29 -20 -26 -27 -17 -24 -25 -23 28 14 -13 11 02 13 15 Treboundophycea esp. -6 -5 14 -13 11 -20 -26 -75 18 26 21 17 24 25 16 21 23 23 23 23 <th< th=""><th>Chlorella vulgaris</th><th>-7</th><th>-1</th><th>-4</th><th>-16</th><th>-8</th><th>28</th><th>23</th><th>22</th><th>19</th><th>-10</th><th>-9</th><th>-11</th><th>13</th><th>-14</th><th>5</th><th>6</th><th>-12</th><th>30</th><th>15</th><th>-18</th><th>-3</th><th>2</th><th>-21</th><th>-29</th><th>-20</th><th>-26</th><th>-27</th><th>-17</th><th>-24</th><th>-25</th></th<>	Chlorella vulgaris	-7	-1	-4	-16	-8	28	23	22	19	-10	-9	-11	13	-14	5	6	-12	30	15	-18	-3	2	-21	-29	-20	-26	-27	-17	-24	-25
Parachlorella kessleri -10 -3 -2 -6 -5 12 -7 -8 -10 -3 -2 5 8 -16 21 -23 -20 -26 -27 -17 -24 -25 -28 12 17 20 23 22 17 24 25 23 24 16 21 23 23 24 17 24 25 23 24 17 24 25 23 24 16 21 23 24 17 24 25 23 24 14 -13 11 20 23 24 17 24 25 23 24 14 13 11 30 3 10 22 13 18 15 16 21 13 12 23 20 26 27 -17 24 25 25 26 21 23 23 26 27 28 20 23 25 26 21 21 23 21 11 21 21 23 23 <th>Chlorella variabilis</th> <th>9</th> <th>U</th> <th>-12</th> <th>5</th> <th>6</th> <th>1</th> <th>-15</th> <th>-30</th> <th>14</th> <th>-13</th> <th>11</th> <th>28</th> <th>23</th> <th>22</th> <th>19</th> <th>25</th> <th>24</th> <th>17</th> <th>21</th> <th>26</th> <th>20</th> <th>23</th> <th>21</th> <th>8</th> <th>16</th> <th>4</th> <th>1</th> <th>2</th> <th>3</th> <th>18</th>	Chlorella variabilis	9	U	-12	5	6	1	-15	-30	14	-13	11	28	23	22	19	25	24	17	21	26	20	23	21	8	16	4	1	2	3	18
Coccomyas subellipsoidea C-169 4 1 12 7 -6 -5 14 -13 11 30 2 3 7 4 1 12 7 24 -15 -16 -10 -18 -16 -10 -18 -16 -10 -16 -10 -18 -10 -18 -16 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10 -10	Parachlorella kessleri	-18	-3	-2	-0	-0	2	3	U.	-30	19	-13	11	20	23	22	19	25	24	1/	21	20	20	23	21	0	4	10	1	10	1
Treboundophyceaesp. To To <thto< th=""> To To T</thto<>	Coccomyxa subellipsoidea C-169	4	1	14	-12	-0	-0	-4	0 0	7	- 10	1	12	-20	-20	-10	20	22	25	-23	-20	27	20	20	20	21	16	22	19	10	10
Mesostigma viride 0 r. N -10 r. -11 -20 <th< th=""><th>Trebouxiophyceae sp.</th><th>0</th><th>7</th><th>14</th><th>-10</th><th>11</th><th>-21</th><th>-20</th><th>-20</th><th>-20</th><th>-27</th><th></th><th>24</th><th>-25</th><th>-15</th><th>2</th><th>20</th><th>23</th><th>25</th><th>-10</th><th>-20</th><th>12</th><th>20</th><th>20</th><th>20</th><th>21</th><th>20</th><th>-22</th><th>22</th><th>10</th><th>10</th></th<>	Trebouxiophyceae sp.	0	7	14	-10	11	-21	-20	-20	-20	-27		24	-25	-15	2	20	23	25	-10	-20	12	20	20	20	21	20	-22	22	10	10
Lhaetosphaeridium gluobosum 42 19 30 10 4 1 16 11 10 14 23 23 73 12 15 18 10 14 24 25	Mesostigma viride	22	10	20	10	1	1	12	-20	12	-21	20	-22	-23	-15	15	3	-	-4	-10	-30	12	0	- 21	-20	-20	20	-23		15	-10
	Chaetosphaeridium globosum Staurastrum punctulatum	-3	-2	30	15	-18	16	9	10	7	-12	-19	22	25	24	17	27	26	20	29	21	-11	13	-14	28	23	8	-1	-4	-6	-5

Table 2: Gene order data set

*where the number denotes genes i.e 1- atpA, 2-atpB, 3-atpE, 4- atpH, 5-psaA, 6-psaB, 7-psbA, 8-psbD, 9-psbE, 10-psbF, 11-psbH, 12 -psbK, 13 -psbN, 14-psbT, 15-rbcL, 16-rps2, 17-rps3, 18rps4, 19-rps7, 20-rps8, 21-rps11, 22-rps12, 23-rps18, 24-rps19, 25-rpl2, 26- rpl14, 27-rpl16, 28rpl20, 29-rpl36, 30-rpoC2 and minus sign denotes gene is present on complementary strand. Genes involved in the common functioning pathway were clustered together as it is evident from previous studies of plants, animal and fungal genomes [3, 21]. A total of 7 gene order clusters are found in our analysis (Table 3).

Saragadam et al RJLBPCS 2019

www.rjlbpcs.com

Gene order	Absent in
rpl2-rps19-rpl22-rps3	Chlorophyta
psbH-psbN-psbT	Euglenophyta
atpB-atpE	Nannochloropsisgaditana(Heterokontophyta),Chalymadomonasreinhardtii,ScenedesmusobliquusandDunaliellasalina(Chlorophyceae)
psaA-psaB	C. reinhardtii, S. obliquus and D. salina (Chlorophyceae), Pseudendoclonium akinetum (Chlorophyta)
psbE-psbF	C. reinhardtii, S. obliquus and D. salina (Chlorophyceae)
rps12-rps7	C. reinhardtii, S. obliquus and D. salina (Chlorophyceae)
rpl36-rps11	<i>C. reinhardtii</i> , <i>S. obliquus</i> and <i>D. salina</i> (Chlorophyceae) and Euglenophyta

Table 3: A total of 7 patterns of gene order are seen among microalgae

Most of the genes are organized into dicistronic clusters. rpl2-rps19-rpl22-rps3 gene order is conserved among all lineages of microalgae except in Chlorophyta (Table 3). rpl2-rps19 gene cluster is present is all the microalgae studied. This indicates the co-transcription of rpl and rps genes. This case is even observed in land plant chloroplasts [22]. From previous studies dicistronic clusters of rpl2-rps19 were detected using RNA gel blots in C. reinhardtii cp DNA clusters of the genes. Cotranscription of ribosomal protein clusters were seen in case of C. reinhardtii chloroplast DNA. This gain the possibility of co-transcription of other functionally related genes [23]. psbH-psbN-psbT is present in all Heterokontophyta, Cryptophyta, Haptophyta, Rhodophyta and Chlorophyta, in Euglenophyta this gene order is absent. Majority of the gene orders are conserved in pairs. atpBatpE present in all microalgae except Nannochloropsis gaditana and Chlorophyceae class, atpAatpH pair is absent in S. obliquus and D. salina and Parachlorella kessleri. psaA-psaB present in all except the case of S. obliquus and D. salina. Genes psbE-psbF are adjacent to each other in all microalgae species but they are dispersed in Chlorophyceae. Genes order rps12-rps7 occurs among all genomes except in Chlorophyceae class and rpl36-rps11 absent in Euglenophyta and Chlorophyceae (Table 3). But in Dinophyta, Heterokontophyta, Haptophyta, Cryptophyta and Rhodophyta the rpl36-rps11 and rps7 and rps12 are present adjacent to each other (Table 3). This type of conservation of gene order in different lineages is considered to be a preserved synteny. *atpB-atpE* encodes the beta and epsilon subunits, respectively, of the membrane component CF1. The *atpB-atpE* genes are cotranscribed as a dicistronic message. A special feature of the *atpB-atpE* genes from several plant species, which is distinct from the other chloroplast atp genes, is

Saragadam et al RJLBPCS 2019 www.rjlbpcs.com Life Science Informatics Publications overlapping of the stop codon of *atpB* with the start codon of *atpE* in the sequence ATGA [24, 25]. atpB-atpE cluster forms the same component in the photosynthetic apparatus. In the thylakoid membranes of chloroplasts and cyanobacteria, the PSI complex mediates light-induced electron transport from plastocyanin or cytochrome c6 to ferredoxin. In cyanobacteria, PSI consists of eleven major polypeptides. Among them, a heterodimer of two integral membrane proteins, psaA and psaB, which form the core of the PSI reaction centre, that harbours nearly 100 antenna chlorophyll (Chl) molecules, and binds the majority of the electron transfer intermediates: P700, a Chl dimer; A0, a monomeric Chl; A1, a phylloquinone and Fx, an iron-sulfur centre. The genes for the two polypeptides are located adjacent to each other in all lineages of microalgae. These genes are present in the large single copy region of the circular plastid genome in higher plants and co-transcribed as a single transcript in rice [26-28]. The *psbE* and *psbF* genes encode the alpha subunit and the beta subunit, respectively, of Cyt b-559 [29]. The alpha and beta subunits are integral thylakoid membrane proteins that form a heme-bridged heterodimer with the heme group bound by conserved His residues located on the transmembrane helices of each subunit [30]. Cyt b-559 is closely associated with oxygen evolving complex Photosystem II reaction centre and found to protect the PSII from photo inhibition [31]. Genes with related function are transcribed as dicistronic or polycistronic transcription unit [32]. From gene order data it is evident that complete rearrangement of all syntenic genes is noticed in Chlorophyceae members. Except for retention of the gene block i.epsbH-psbN-psbT (Figure 1). High degree of gene order variation is seen in the case of Chlorophyceae members (Figure 1). Complete gene order conservation is seen between genus groups of Phaeocystis antarctica and Phaeocystis globosa, Euglena gracilis and Euglena viridis where no gene arrangements is noticed (Figure 1). Such conservation is lost between genus group of Thalassiosira pseudonana and Thalassiosira oceanica were there is no correlation of gene order (Figure 1).

Saraga	dam et al	I RJLB	PCS 201	9

psaB

rps2

rpoC2

rps18

atpB

atpE

psbK

Hdsq

psbN

psbT

Odaq

atpA

atpH

rbcL

rpl20

PsbA

rps4

rp12

rps19

rps3

rp116

rpl14

rps8

rps11 rpl36

rps12

rps7

Odontella sinensis

rps7

rps12

rps11

rpl36

rps8

rpl14

rp116

rps3

rps19

rpl2

Tps4

psbA

rps4

psbK

atpE

atpB

rps18

rpoC2

Tps2

rpl20

hdsq

Ndsq

psbT

atpA

atpH

<u>hcl</u>

psaB

Thalassiosira pseudonana

rps7

rps12

rps11

rpl36

rps8

rpl14

rpl16

rps3

rps19

rp12

rps4

PsbA

PsaA

psaB

atpA

atpH

rbcl

rpl20

Hdsq

Ndsq

psbT

psbK

Ddsq

atpE

rps18 atpB

TpoC2

rps2

Thalassiosira oceanica

rpl14

rps8

rp136

rps11

rps12

rps7

rp|20

atpE

atpB **Fistulifera**

rps18

psbA

rp120

psaB

psaA

rps4

psbK

Ddsq

rbcL

psbK

Ddaq

rps7

rps12

rps11

rpl36

rps8

rpl14

rp116

rps3

rps19

rp12

rps4

psbA

rpl20

atpA

atpH

rbcL

rps2

TpoC2

rps18

atpB

atpE

psbT

Ndsq

Hdsq

psbK

Ddaq

Aeso

osaB

Phaeodactylum tricornutum

rps7

rps12

rps11

rpl36

rps8

rp114

rp116

rps3

rps19

rpl2

rps4

psbA

rpl20

atpA

atpH

막

rps2

TpoC2

rps18

atpB

atpE

psbT

psbN

Hdsq

psbK

Ddaq

Durinskia baltica

psaB

psbA

rps4

rps7

rps12

rps11

rpl36

rps8

rpl14

rpl16

rps3

rps19

rpl2

rpl20

atpA

atpH

Pcl

rps2

rpoC2

rps18

atpB

atpE

psbT

Ndsq

Hdsq

psbK

psaA psaB psbD

Kryptoperidinium foliaceum

www.rjlbpcs.com rpoC2 psbT **Ndsq** rps2 **Hdsq** atpA atpH rpl2 rps19 눧 psbT rps3 rp116 **Ndsq** rpl14 Hdsd rps8 psbK rpl36 Dabo <u>rps11</u> rps12 rps7 atpB psbA atpE rps4 atpH atpA rpl2 rps19 rps18 rps3 rpoC2 rp116 rps2

Het	erosić	a ma	ikashi	M						-		-																	
sbK	rpl20	rpoC	2 rps	2 atpl	H atp	A psb		2 rps1	9 rps3	rpl16	5 rpl1	4 rps8	rpl36	rps11	l rps1.	2 rps7	Tdsq 1	Ndsq	Hdad	psbf	psbE	psaA	psaB	rbcL	rps4	rps18	atpB	atpE	
Aur	eococ	o sno:	nopha	ıgeffer	suar	8																							
Dda	rpl2	rps19	rpl16	5 rps3	rpl14	4 rps8	rpl3	6 rps1	1 rps1	2 rps	7 psb	T psbl	N psbi	H psbF	psbE	psaA	psaB	rbcL	TpoC	rps4	rps18	rps2	atpH	atpA	atpE	atpB	psbK	rpl20	a
Aur	unoa.	ibra lo	aunbc	nsis																3									
psaA	psaB	bsbk	Adsq 2	H psb/	dsg N	T rps7	rps.	12 rps	11 rpla	<u>36 rps</u>	5 <u>8</u> rpl1	<u>4 rpl:</u>	<u>16</u> <u>rps</u>	<u>3 rps1</u>	<u>9</u> rpl2	Ddsq	rpl20	psbE	psbF	rps4	psbA	<u>hc</u>	atpE	<u>ps18</u>	rpoC2	rps2	atpH	atpA	(0)
Na	nnoch	ilorop	sis ga	dītana																									
rps18	3 psb/	A rpl2	2 rps1	9 rps3	3 rpl1	.6 rpl1	4 rps	s8 rpl3	6 rps1	11 rps	12 rp.	s7 ps	A psa	B atpl	E atpB	psbK	psbT	psbN	Hdsq	psbE	psbF	Ddsq	atpA	atpH	<u>rps2</u>	rps4	TPOC2	rbcL	8
ρμα	Jeocy	stis aı	ntarcti	ica																									
										100	No.				and the second sec				1			1000		1 100 V	100		NOTES .	20 02	

TPOC2 rbcl rpl20 rps7 rps12 rps11 psbD atpA atpH rps2 rps4 rpl36 rps8 rp|14 rp116 rps3 rps19 rp12 Ddaq psbK atpA atpB atpH atpE Tps2 rps4 rpoC2 rps7 hdsq rps12 psbN rps11 psbT rpl36 psbK rps8 atpB rpl16 rpl14 atpE rps18 <u>rp|20</u> rpl2 rps19 **psbA** rps1 Phaeocystis antarcti Phaeocystis globosa Emiliania huxleyi rpl2 rbcL psbA **psbA** psaB rps18 rps18 psaA

Life Science Informatics Publications

atpB

atpE

psaB

psaA

psbA

rps18

atpH

atpA

rpl20

rpoC2

rps2

Ddaq

rpl36

rps11

rps12

rps7

bsbK

rps4

Hdsq

psbN

psbT

rbcL

rps8

rpl14

rpl2 rps19 rps3 rpl16

Pavlova lutheri

_						_			n i	_										_					
Hdsq		Hdad		psaA		psbE		rpl36		rpoC2		rpoC2		rps7		atpA		rps2		psbA		rps2		<u>rbcl</u>	
Ndsq		Ndsq		psaB	1	psbf		rps11		rps11		rps11		rps12		atpH		rpoC2		rbcL		rps18		psaB	
psbT		psbT		bsbK		atpA		psbK		rps4		rps4		rpl20		rps2		atpH		psbK		<u>rpl20</u>	1	psaA	
rps7		rps7		<u>rpl20</u>		atpH		Ddsq		psbT		psbT		rps4		psbT		atpA		rpl20		psaB		psbT	
rps12		rps12		rpl2		rps2		atpE		Hdsq		Hdsq		psaA		psbH		rps7		psaA		psaA		Ndsq	
rps11		rps11		rps19		rpoC2		atpB		Ndsq		Ndsq		psaB		psbN		rps12		psaB		rps4	10.	Hdsq	
rpl36		rpl36		rps3		rps19		PsbA		atpB		atpB		psbE		rpoC2		<u>120</u>		atpA		psbA		<u>rp120</u>	
rps8		rps8		<u>rpl16</u>		atpB		rbcL		atpE		atpE		psbF		<u>rps4</u>		rps2		atpH		psbT		rps4	
rpl14		rpl14		rpl14		atpE		TPoC2		<u>rbcl</u>		I pc		rpl2		<u>rps11</u>		psaA		rps11		Ndsq		<u>rps18</u>	
rpl16		rpl16		rps8		rps7		rps2		rps18		rps18		rps19		rps18		psaB		rps8		Hdsq		psbE	
rps3		rps3		<u>rp136</u>		rps12		atpH		atpA		atpA		rps3		psbF		atpE		rpl14		Ddsq		psbf	
rps19		rps19		<u>rps11</u>		rps11		atpA		atpH		atpH		rpl16		psbE		atpB		rpl16		rps11		psbK	
rpl2		rpl2		rps12		rp136		rps4		rps2		rps2		rpl14		psaB		밀		rps3		rp136		rps7	
psaA		psaA		<u>rps7</u>		rps8		rps12		rp136		rpl36		rps8		psaA		<u>rpl2</u>		rps19		rps8		rps12	
psaB		psaB		rbcL		rpl14		rps7		rps8		rps8		rpl36	0	rps7		rps19		rp12		rpl14		Ddaq	
bsbK		psbK		rps4		rpl16		rps18		rpl14		rpl14		TDS2		rpl12		rps3		rps2		rpl16		rps2	
<u>rpl20</u>		rpl20		psbT		rps3		rp12		rpl16		rpl16		atpH		rp120		rp116		rpoC2		rps3		rpoC2	
atpE		atpE		Ndaq		rps19		rps19		rps3		rps3		atpA		Ddsq		rpl14		rps7		rps19		<u>atpH</u>	
atpB		atpB		Hdad		rpl2		rps3		rps19		rps19		<u>rps18</u>		psbK		<u>irps8</u>		rps12		rp12		atpA	
rps18		rps18		Adaq		rpl20		<u>rpl16</u>		rpl2		rpl2		무다		psbA		rp136		rps18		atpA	2	rpl2	
TpoC2		TPOC2		psbE		psbK		rpl14		psbF		psbf		3 atpE		rp136		L rps4		rps4		atpH		rps19	
rps2		rps2		psbf		psaB		rps8		psbE		psbE		<u>N</u> atp		rps8		rps1.		Ndsq		rps7		rps3	
atpH		atpH		atpA	2	psaA		psbF		psaB		psaB		H psbl		rpl14		rps18		<u>psbH</u>		rps12		rp116	
atpA		atpA		atpH	5	Hdsq	ao	psbE		psaA		psaA		E psbł		rpl16		<u> Yqsq</u>		psbE		psbK		rpl14	
<u>psbE</u>		psbE		rps2	5	Ndsq	train 1	Hdsq		<u>rp120</u>		<u>rp120</u>		2 psb	g	rps3		psbE		psbF		pisbiF		rps8	
psbF		psbF		rpoC2	r	psbT	olae si	psbT		<u>rps12</u>		rps12		TpoC	Imatic	rps19	2	psbF	e	atpB		psbE	ij	<u>rp136</u>	0
rbcl.		rbcL	lina	<u>rps18</u>	arium	Ddaq	n mer	Ndaq	adoxi	rps7		<u>rps7</u>		rps11	aenig	rpl2	nnasti	psbT	arkea	atpE	iuri	atpE	ospac	<u>rps11</u>	livace
rps4	theta	rps4	nas sa	atpE	ı caldı	PsbA	chyzoi	rp120	ra pai	psbK	racilis	psbK	iridis	psbK	phina	rbcL	la gyn	Ndsq	d spue	<u>rpl36</u>	cus ta	atpB	ud sn:	atpE	mis o
psbA	lardia	psbA	lomot	atpB	nidiun	ЪсL	nidios	psaB	oydou	PsbA	ena g	PsbA	ena v	PsbA	omor	atpE	eptiel	Hdag	Imimo	<u>psbT</u>	20203.	rpoC2	30000	atpB	hrose
Odaq	Guil	Ddaq	Rho	Ddaq	Cyar	rps4	Cyar	psaA	Cyar	Ddaq	Eugl	Ddsq	Eug	Ddsq	Mor	atpB	Eutr	psbA	Руга	DabD	Ostr	rbcl	Руа	psbA	Nep

Saragadam et al RJLBPCS 2019

www.rjlbpcs.com

Life Science Informatics Publications

Ddsq

rps12

atpB

psbA

rps11

atpH

atpA

<u>rbcl</u>

psaB

Tood

rps3

<u>rp120</u>

rps2

psbT

Ndsq

Hdsq

rps7

atpE

rps4

psaA

rps8

psbK rpl20 rpl36 rpl2 rps19 rpl16 rpl14 Chlamydomonas reinhardtii

psbK

psbA

rps18

rpl2 rps19 rps3 rpl16 rpl14 rps8

psbT psbN

psbK

PsbH

rps11

rps4

rps2

atpB

atpE

atpH

atpA

rpl36

TpoC2

rbcL

<u>rpl20</u>

psaA

<u>rps12</u>

rps7

Ddsq

Monomastix sp. OKE-1



Figure 1: Gene order of the conserved protein coding chloroplast genes in microalgae. Gene order is based on GenBank sequences of microalgae.

© 2019 Life Science Informatics Publication All rights reserved Peer review under responsibility of Life Science Informatics Publications 2019 March – April RJLBPCS 5(2) Page No.211

Saragadam et al RJLBPCS 2019

www.rjlbpcs.com

3. Phylogenetic inference from gene order data

Maximum likelihood tree is inferred from gene order data set using MLGO with 1000 bootstrap replicates (Figure 2). But the gene order failed to resolve phylogeny relationship among the entire microalgae except in the case of Euglenophyta group (Figure 2).



Figure 2: Gene order Maximum likelihood tree generated by using MLGO (Maximum Likelihood for gene order analysis). Numbers at nodes denote bootstrap values for 100 replicates.

Monophyly in Euglenophyta is evidently supporting that the strong conservation in its gene order. A well supported monophyletic clade is observed among the Chlorophyceae class. Organisms belonging to same genus i.e *Euglena gracilis and Euglena viridis, Phaeocystis antarctica* and

Saragadam et al RJLBPCS 2019 www.rjlbpcs.com Life Science Informatics Publications *Phaeocystis globosa* were clustered together in the same clade (Figure 2). As these microalgae share the same gene order, genome rearrangements are less prone to occur in closely related species [33].

4. Synteny

To assess the extent of chloroplast genome rearrangements more closely, genomes were aligned using the Progressive Mauve aligner version 2.1. 55 syntenic regions, which are free from obvious genome rearrangement, are apparent in the synteny map (Figure 3).







Saragadam et al RJLBPCS 2019

www.rjlbpcs.com

Life Science Informatics Publications



Figure 3: Alignments showing syntenic rearrangements in the 43 chloroplast genomes of microalgae using Progressive Mauve algorithm. Corresponding colour boxes are locally collinear blocks (LCBs). The height of colored lines inside LCB represents the similarity of the blocks. White boxes are protein coding genes. The LCB present below the axis indicates inversion. Red boxes are rRNA genes, green boxes are tRNA genes, large Pink boxes are repeat regions. LCB1-Chll, trnL(caa); LCB2-psbD; LCB3-psbD, psbC, orf116; LCB4-orf31, psbC, orf462; LCB5-trnL(uaa); LCB6-psbA; LCB7-psbA; LCB8-psbA, ycf65, trnR(ucu); LCB9-psbK, trnT(ugc), trnG(gcc), trnM(cau); LCB10-psaI, trnY(gua), rps4, rps11, trnL(uag), trnR(acg), trnV(uac), rpoC2; LCB11-rpoC2; LCB12-rpoC2,rpoC1; LCB13-rpoC1; LCB14-rpoB; LCB15-rpoB, trnH(gug), trnM(cau), trnW(cca), trnE(uuc), trnG(ucc); LCB16-psbB, psbT, psbH; LCB17-psbN, petB; LCB18-petB; LCB19-atpB; LCB20-atpB; LCB21-atpB, atpE; LCB22-atpE, rbcL; LCB23-rbcL; LCB24-rbcL, rpl32, psaC; LCB25-psaC; LCB26-rps9, rpl12,

Saragadam et al RJLBPCS 2019 www.rjlbpcs.com Life Science Informatics Publications psbZ, trnS(uga), trnP(ugg); LCB27-psaJ, rps18; LCB28-rps18,atpA; LCB29-atpF, atpH; LCB30-atpI, rps2; LCB31-rps2, trnC(gca), trnF(gaa); LCB32-rp14, trnI(cau), rpl36, rps8, rpl5; LCB33-rpl16; LCB34-rpl16; LCB35-roaA, orf240, rps3; LCB36-rps3, rpl22, rps19, rpl2; LCB37-rpl23, psbJ, psbL, psbF, psbE; LCB38-psaB; LCB39-psaB; LCB40-psaA; LCB41-psaA; LCB42-trnK(uuu), petG; LCB43-trnD(guc), psbI; LCB44-rps12,rps7; LCB45-tufA; LCB46-ycf4, trnQ(uug), trnS(gcu); LCB47-rrl; LCB48-rrl; LCB49-rrl; LCB50-rrl; LCB51-rrl; LCB52-trnA(ugc),trnI(gac); LCB53-rrs; LCB54-rrs; LCB55-rrs present in the reference genome.

Despite an overall highly conserved gene content between the phylum and class, the genomes have multiple rearrangements as indicated in the Mauve alignment (Figure 3). Out of 55 syntenic LCB clusters present in the reference set genome, only 25 LCBs (underlined LCBs in figure 3 legend) are present in common with some exceptions. Most of the LCBs are missing among Chlorophyceae class. The chloroplast genomes of microalgae were poorly collinear, whereas collinear alignment in the structure was observed between members of the phylum (Figure 3). The arrangements of tRNA genes are highly variable among microalgae and pattern of synteny is completely absent. Conserved clusters of *rpoC2-rpoC1*, *atpB-atpE* (absent in Chlorophyceae), *rps18-atpA* (absent in *E. gracilis*, *E. viridis*), *atpG-atpH*(*atpF-atpH* present in Euglenophyta group), *rps12-rps7* (absent in Chlorophyceae) genes present in the genomes of microalgae. Moreover gene order scrambled for LCB clusters *psbB-psbT-psbH/psbN*, *rps3-rpl22-rps19-rpl2* and *rpl23-psbJ-psbL-psbF-psbE*.

4. CONCLUSION

From this analysis, it indicates that Chlorophyceae and Euglenophyceae members have highly disarranged and scrambled plastid genome when compared with all other microalgal species. The chloroplast gene order of the Chlorophycean group of algae is unique and differs from those of other microalgal species. From the results, it is evident that plastid gene order is not applicable in inferring phylogenetic relationships among microalgae except at species level.

ACKNOWLEDGEMENT

We specially thank to Mr. Ravi Kanth Reddy Sathi for his support in sequence file retrieval. Funding support provided by CSIR UGC.

CONFLICT OF INTEREST

Authors have no any conflict of interest.

- 1. Zeng X, Nesbitt MJ, Pei J, Wang K, Vergara IA, Chen N, OrthoCluster: a new tool for mining synteny blocks and applications in comparative genomics. In the Proceedings of the 11th international conference on Extending database technology: Advances in database technology 2008, pp. 656-667.
- Wolf YI, Rogozin IB, Kondrashov AS, Koonin EV. Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome research. 2001 Mar 1; 11(3):356-72.
- 3. Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of eukaryotic gene order. Nature Reviews Genetics. 2004 Apr; 5(4):299.
- 4. Downie SR, Palmer JD. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. Edn, Soltis PS, Soltis DE and Doyle JJ, Springer, Boston, MA, 1992, pp. 14-35.
- 5. Palmer JD. Mitochondrial DNA in plant systematics: applications and limitations. Edn, Soltis PS, Soltis DE and Doyle JJ, New York, USA, 1992, pp. 36-49.
- Raubeson LA, Jansen RK. Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science. 1992 Mar 27; 255(5052):1697-9.
- Olmstead RG, Palmer JD. Chloroplast DNA systematics: a review of methods and data analysis. American journal of botany. 1994 Sep; 81(9):1205-24.
- Belda E, Moya A, Silva FJ. Genome rearrangement distances and gene order phylogeny in γproteobacteria. Molecular Biology and Evolution. 2005 Mar 16; 22(6):1456-67.
- Luo H, Shi J, Arndt W, Tang J, Friedman R. Gene order phylogeny of the genus Prochlorococcus. PLoS One. 2008 Dec 3; 3(12):e3837.
- Luo H, Sun Z, Arndt W, Shi J, Friedman R, Tang J. Gene order phylogeny and the evolution of methanogens. PLoS One. 2009 Jun 29; 4(6):e6069.
- 11. Blanchette M, Kunisawa T, Sankoff D. Gene order breakpoint evidence in animal mitochondrial phylogeny. Journal of Molecular Evolution. 1999 Aug 1; 49(2):193-203.
- Jansen RK, Raubeson LA, Boore JL, Chumley TW, Haberle RC, Wyman SK, Alverson AJ, Peery R, Herman SJ, Fourcade HM, Kuehl JV. Methods for obtaining and analyzing whole chloroplast genome sequences. In Methods in enzymology. 2005, 395:348-384.
- Zimmer EA, Roalson E. Molecular Evolution, Producing the Biochemical Data. Elsevier; 2005 Jun 15.
- 14. Maul JE, Lilly JW, Cui L, Miller W, Harris EH, Stern DB. The Chlamydomonas reinhardtii plastid chromosome: islands of genes in a sea of repeats. The Plant Cell. 2002 Nov 1; 14(11):2659-79.
- Poyatos JF, Hurst LD. The determinants of gene order conservation in yeasts. Genome biology. 2007 Nov 5; 8(11):R233.

Saragadam et al RJLBPCS 2019 www.rjlbpcs.com Life Science Informatics Publications
16. Hu F, Lin Y, Tang J. MLGO: phylogeny reconstruction and ancestral inference from gene-order data. BMC bioinformatics. 2014 Nov 8; 15(1):354.

- 17. Darling AE, Mau B, Perna NT. ProgressiveMauve:multiple genome alignment with gene gain, lossand rearrangement. PloS one. 2010 Jun 25; 5(6):e11147.
- Robbens S, Derelle E, Ferraz C, Wuyts J, Moreau H, Van de Peer Y. The complete chloroplast and mitochondrial DNA sequence of Ostreococcus tauri: organelle genomes of the smallest eukaryote are examples of compaction. Molecular biology and evolution. 2007 Jan 23; 24(4):956-68.
- 19. Gray MW, Lang BF, Burger G. Mitochondria of protists. Annu. Rev. Genet. 2004 Dec 15; 38:477-524.
- 20. Green BR. Chloroplast genomes of photosynthetic eukaryotes. The plant journal. 2011 Apr; 66(1):34-44.
- 21. Lee JM, Sonnhammer EL. Genomic gene clustering analysis of pathways in eukaryotes. Genome research. 2003 May 1; 13(5):875-82.
- 22. Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K. RNA editing in hornwort chloroplasts makes more than half the genes functional. Nucleic Acids Research. 2003 May 1; 31(9):2417-23.
- 23. Cui L, Leebens-Mack J, Wang LS, Tang J, Rymarquis L, Stern DB. Adaptive evolution of chloroplast genome structure inferred using a parametric bootstrap approach. BMC evolutionary biology. 2006 Feb 9; 6(1):13.
- 24. Zurawski G, Bottomley W, Whitfeld PR. Structures of the genes for the β and ε subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. Proceedings of the National Academy of Sciences. 1982 Oct 1; 79(20):6260-4.
- 25. Gatenby AA, Rothstein SJ, Nomura M. Translational coupling of the maize chloroplast atpB and atpE genes. Proceedings of the National Academy of Sciences. 1989 Jun 1; 86(11):4066-70.
- 26. Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. The EMBO journal. 1986 Sep; 5(9):2043-9.
- 27. Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ. The complete sequence of the rice (Oryza sativa) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Molecular and General Genetics MGG. 1989 Jun 1; 217(2-3):185-94.
- Chen SC, Cheng MC, Chung KR, Yu NJ, Chen MC. Expression of the rice chloroplast psaApsaB-rps14 gene cluster. Plant Science. 1992 Jan 1; 81(1):93-102.

Saragadam et al RJLBPCS 2019 www.rjlbpcs.com Life Science Informatics Publications
29. Andersson BE, Styring ST. Photosystem II: Molecular organization, function, and acclimation. Curr. Top. Bioenerg. 1991 Jan 1; 16:1-81.

- 30. Cramer WA, Tae GS, Furbacher PN, Böttger M. The enigmatic cytochrome b-559 of oxygenic photosynthesis. Physiologia plantarum. 1993 Aug; 88(4):705-11.
- Thompson LK & Brudvig GW. Cytochrome b-559 may function to protect photosystem II from photoinhibition. *Biochemistry* 1988 Sep 6; 27(18):6653-6658.
- 32. Sugita M, Sugiura M. Regulation of gene expression in chloroplasts of higher plants. Plant molecular biology. 1996 Oct 1; 32(1-2):315-26.
- 33. Boore JL. Animal mitochondrial genomes. Nucleic acids research. 1999 Apr 1; 27(8):1767-80.