



Genome-wide comparative analysis of codon usage bias and codon context patterns among cyanobacterial genomes



Ratna Prabha^a, Dhananjaya P. Singh^{a,*}, Swati Sinha^{b,d}, Khurshid Ahmad^a, Anil Rai^c

^a ICAR-National Bureau of Agriculturally Important Microorganisms, Indian Council of Agricultural Research, Kushmaur, Maunath Bhanjan 275103, India

^b Centre for Computational Biology and Bioinformatics, School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, India

^c ICAR-Indian Agricultural Statistical Research Institute, Indian Council of Agricultural Research, Pusa, New Delhi 110012, India

^d Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01 Matrix, Singapore City 138671, Singapore

ARTICLE INFO

Article history:

Received 14 June 2016

Received in revised form 11 September 2016

Accepted 3 October 2016

Available online 10 October 2016

Keywords:

Cyanobacteria

Codon bias

Codon selection pattern

Codon context

GC composition bias

ABSTRACT

With the increasing accumulation of genomic sequence information of prokaryotes, the study of codon usage bias has gained renewed attention. The purpose of this study was to examine codon selection pattern within and across cyanobacterial species belonging to diverse taxonomic orders and habitats. We performed detailed comparative analysis of cyanobacterial genomes with respect to codon bias. Our analysis reflects that in cyanobacterial genomes, A- and/or T-ending codons were used predominantly in the genes whereas G- and/or C-ending codons were largely avoided. Variation in the codon context usage of cyanobacterial genes corresponded to the clustering of cyanobacteria as per their GC content. Analysis of codon adaptation index (CAI) and synonymous codon usage order (SCUO) revealed that majority of genes are associated with low codon bias. Codon selection pattern in cyanobacterial genomes reflected compositional constraints as major influencing factor. It is also identified that although, mutational constraint may play some role in affecting codon usage bias in cyanobacteria, compositional constraint in terms of genomic GC composition coupled with environmental factors affected codon selection pattern in cyanobacterial genomes.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Codon usage bias (codon bias) is a phenomenon of usage of specific codons more frequently than other synonymous codons by the organisms during the translation of genes to proteins (Salim and Cavalcanti, 2008; Baeza et al., 2015). It represents balance between mutational forces and translational selection that leads to translational efficiency of genes (Qiu et al., 2011; Behura and Severson, 2012). Non-random usage of synonymous codons facilitates protein-coding sequences to maintain information at the genomic level (Gilchrist et al., 2009). Evidences suggest that selection pressure shapes codon usage bias (Chen et al., 2004) and is related to gene expression where high expressed genes have strong codon usage bias in comparison to low-expressed genes (Bulmer, 1991; Carbone et al., 2003). Highly expressed genes possess a preference towards codons that facilitate accurate and efficient translation of genes into proteins (Bulmer, 1988; Sharp, 1989). Since majority of genes are not highly expressed, the codon usage is generally influenced by mutations rather than selection (Henry, 2007).

With rapid progress in whole genome sequencing, analysis of codon usage bias at the genome level, rather than on a single or a set of gene,

has gained attention. Genome-wide investigations on the variation of codon and codon context bias are important for understanding of functional evolution of genome within and between species. Apart from the codon usage, the codon context sequences (i.e. sequential pair of codons) that correspond to A- and P- binding sites of ribosome during translational process are also important features to define translational selection (Behura and Severson, 2012). Nucleotide substitutions and genomic GC content are considered as the principal determinants of codon usage in prokaryotes (Chen et al., 2004). Other biological factors like GC content, amino acid composition, translation, gene expression level, gene length, tRNA abundance, protein structure and mutation frequency vary in accordance with codon usage bias (Bernardi and Bernardi, 1986; Bains, 1987; Lobry and Gautier, 1994; Karlin and Mrazek, 1996; D'Onofrio et al., 2002; Ma et al., 2002; Gu et al., 2004; Wan et al., 2004; Doherty and McInerney, 2013). Factors like genetic recombination, gene conversion, mutational bias, intron splicing, DNA packaging and selection also modulate translational accuracy as well as efficiency of gene products (Akashi, 2003; Musto et al., 2003; Chen et al., 2004; Hassan et al., 2009; Lin et al., 2006; Warnecke and Hurst, 2007).

Cyanobacteria also known as blue-green algae are one of the eleven main eubacterial phyla. It corresponds to a tremendously diverse group of prokaryotes, owing to variation in their morphology, physiology and developmental features (Beck et al., 2012). Cyanobacteria evolve from ancient photosynthetic prokaryotes, though they exhibit visible

* Corresponding author.

E-mail address: dpsfarm@rediffmail.com (D.P. Singh).

distinctions in their features and physiology (Larsson et al., 2011). They are phototrophic microbes but contrary to rest of the phototrophic bacteria, they utilize water as an electron-donor and thus discharge oxygen (Beck et al., 2012). Cyanobacteria acquire peculiar role in oxygen evolution as they are accountable for oxygenic photosynthesis and therefore changing the atmosphere of earth. Cyanobacteria also exhibits specific role in the global carbon and nitrogen biogeochemical cycles (Larsson et al., 2011). Cyanobacteria are the most primitive photosynthetic prokaryotic organisms representing five morphologically and taxonomically distinct groups (Rippka et al., 1979; Mareš et al., 2013). Evolutionary history coupled with exceptional diversity of cyanobacteria make these organisms suitable for genome-wide studies on their codon usage. Analysis of codon usage in sequenced genomes of cyanobacteria occupying different natural habitats such as marine, fresh water, hot water, terrestrial and hard rock is important as it can provide insights into their adaptation to the environment and evolutionary diversification. One of the objectives of this study is to investigate the codon usage in the genomes of different cyanobacterial species that belong to diverse habitats and present different physiological and morphological characteristics. Although some studies have been conducted on the codon usage in certain closely related cyanobacterial genomes (Campbell and Gowri, 1990; Prabha et al., 2012; Yu et al., 2012; Xu et al., 2013), a comprehensive genome-wide analysis representing major taxonomic groups of cyanobacteria is lacking. We analyzed codon usage bias and codon

context patterns among 41 sequenced cyanobacterial genomes representing five different taxonomic orders. Since cyanobacteria possess wide diversity in terms of their habitat, physiology, taxonomy and evolution, this study aims to provide critical insights in to codon usage bias and codon context pattern and correlate this with their adaptation to diverse habitats. Genome-wide investigations of codon bias patterns, their causes, consequences and identification of forces that shape environmental adaptation and genomic evolution are of quite importance with reference to cyanobacterial genomes because these organisms hold a long history of biological diversity and evolution.

2. Materials and methods

2.1. Complete genome sequences

Complete genome sequence of 41 cyanobacterial strains were obtained from NCBI Genome database (Table 1). These genomes belonged to 5 different taxonomic orders i.e. Chroococcales (22 strains), Prochlorales (12 strains), Nostocales (4 strains), Oscillatoriales (2 strains) and Gloeobacterales (1 strain). In order to minimize sampling error and avoid potential length bias, genes with at least 80 codons with correct initiation and termination codons were used for the calculation of codon usage indices (Eyre-Walker, 1996).

Table 1
Details and genomic features of cyanobacterial genomes with Nc and GC3s values.

Order	Organisms	Abbreviation for organisms	Habitats	Size (Mb)	GC content (%)	No. of genes	Nc ^a	GC3s ^b
Chroococcales	<i>Acaryochloris marina</i> MBIC11017	Am-MBIC11017	Marine	8.36	47	8571	54.91	0.467
	<i>Cyanothece</i> sp. ATCC 51142	Cs-ATCC51142	Marine	5.46	37.9	5365	46.73	0.296
	<i>Cyanothece</i> sp. PCC 7424	Cs-PCC7424	Freshwater	6.56	38.5	5942	47.29	0.297
	<i>Cyanothece</i> sp. PCC 7425	Cs-PCC7425	Freshwater	5.78	50.6	5507	51.44	0.535
	<i>Cyanothece</i> sp. PCC 7822	Cs-PCC7822	Freshwater	7.84	40.1	7042	49.97	0.325
	<i>Cyanothece</i> sp. PCC 8801	Cs-PCC8801	Freshwater	4.79	39.8	4619	48.76	0.324
	<i>Cyanothece</i> sp. PCC 8802	Cs-PCC8802	Freshwater	4.80	39.8	4700	48.75	0.326
	<i>Microcystis aeruginosa</i> NIES-843	Ma-NIES-843	Freshwater	5.84	42.3	6364	52.48	0.382
	<i>Synechococcus elongatus</i> PCC 6301	Se-PCC6301	Freshwater	2.7	55.5	2582	51.17	0.602
	<i>Synechococcus elongatus</i> PCC 7942	Se-PCC7942	Freshwater	2.75	55.4	2715	51.44	0.596
	<i>Synechococcus</i> sp. CC9311	Ss-CC9311	Marine	2.60	52.4	2944	55.17	0.537
	<i>Synechococcus</i> sp. CC9605	Ss-CC9605	Marine	2.51	59.2	2756	44.66	0.71
	<i>Synechococcus</i> sp. CC9902	Ss-CC9902	Marine	2.23	54.2	2357	54.07	0.567
	<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	Ss-JA-2-3B'a(2-13)	Hot spring	3.05	58.5	2942	45.88	0.694
	<i>Synechococcus</i> sp. JA-3-3Ab	Ss-JA-3-3Ab	Hot spring	2.93	60.2	2897	43.66	0.736
	<i>Synechococcus</i> sp. PCC 7002	Ss-PCC7002	Marine	3.40	49.2	3238	51.2	0.525
	<i>Synechococcus</i> sp. RCC307	Ss-RCC307	Marine	4.42	60.8	2582	42.61	0.717
	<i>Synechococcus</i> sp. WH 7803	Ss-WH7803	Marine	4.77	60.2	2586	45.84	0.694
	<i>Synechococcus</i> sp. WH 8102	Ss-WH8102	Marine	4.83	59.4	2581	45.73	0.698
	<i>Synechocystis</i> sp. PCC 6803	Sy-PCC6803	Freshwater	3.95	47.4	3625	50.86	0.495
	<i>Thermosynechococcus elongatus</i> BP-1	Te-BP-1	Hot spring	2.59	53.9	2525	50.23	0.572
	<i>cyano bacterium</i> UCYN-A	C-UCYNA	Marine	1.4	31.1	1241	40.97	0.173
Gloeobacterales	<i>Gloeobacter violaceus</i> PCC 7421	Gv-PCC7421	Rock	4.66	62	4482	40.47	0.785
Nostocales	<i>Anabaena variabilis</i> ATCC 29413	Av-ATCC29413	Multiple	7.10	41.4	5813	52.44	0.341
	' <i>Nostoc azollae</i> ' 0708	Na-0708	Multiple	5.53	38.3	5380	49.31	0.29
	<i>Nostoc punctiforme</i> PCC 73102	Np-PCC73102	Freshwater	9.06	41.4	7164	53.53	0.346
	<i>Nostoc</i> sp. PCC 7120	Ns-PCC7120	Multiple	7.2	41.3	6213	52.63	0.348
Oscillatoriales	<i>Arthrospira platensis</i> NIES-39	Ap-NIES-39	Freshwater	6.8	44.3	6676	54	0.395
	<i>Trichodesmium erythraeum</i> IMS101	Te-IMS101	Marine	7.75	34.1	5126	46.09	0.235
Prochlorales	<i>Prochlorococcus marinus</i> str. AS9601	Pm-AS9601	Marine	1.67	31.3	1965	40.18	0.183
	<i>Prochlorococcus marinus</i> str. MIT 9211	Pm-MIT9211	Marine	1.7	38	1900	47.17	0.249
	<i>Prochlorococcus marinus</i> str. MIT 9215	Pm-MIT9215	Marine	1.74	31.1	2054	40.12	0.182
	<i>Prochlorococcus marinus</i> str. MIT 9301	Pm-MIT9301	Marine	1.64	31.3	1962	40.22	0.184
	<i>Prochlorococcus marinus</i> str. MIT 9303	Pm-MIT9303	Marine	2.7	50	3136	55.54	0.508
	<i>Prochlorococcus marinus</i> str. MIT 9312	Pm-MIT9312	Marine	1.70	31.2	1856	40.09	0.183
	<i>Prochlorococcus marinus</i> str. MIT 9313	Pm-MIT9313	Marine	2.4	50.7	2330	55.69	0.502
	<i>Prochlorococcus marinus</i> str. MIT 9515	Pm-MIT9515	Marine	1.7	30.8	1964	39.95	0.181
	<i>Prochlorococcus marinus</i> str. NATL1A	Pm-NATL1A	Marine	1.86	35	2250	43.85	0.217
	<i>Prochlorococcus marinus</i> str. NATL2A	Pm-NATL2A	Marine	1.84	35.1	2228	43.94	0.218
	<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375	Pm-CCMP1375	Marine	1.75	36.4	1930	45.07	0.222
	<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMP1986	Pm-CCMP1986	Marine	1.66	30.8	1762	39.52	0.175

^a Effective number of codons.

^b GC content at third position of codon.

Download English Version:

<https://daneshyari.com/en/article/5518248>

Download Persian Version:

<https://daneshyari.com/article/5518248>

[Daneshyari.com](https://daneshyari.com)