Vol.62: e19170747, 2019
http://dx.doi.org/10.1590/1678-4324-2019170747
ISSN 1678-4324 Online Edition

Article – Biological and Applied Sciences

Genome-Wide Assessment of Putative Superoxide Dismutases in Unicellular and Filamentous Cyanobacteria

Rajesh Prajapati[1]
https://orcid.org/0000-0002-7013-6312

Shivam Yadav[2]
https://orcid.org/0000-0001-9746-2962

Sonali Mitra[1]
https://orcid.org/0000-0001-8945-3062

Priya Rai[1]
https://orcid.org/0000-0001-6314-2354

Rajeev Mishra[1]
https://orcid.org/0000-0001-8559-5258

Neelam Atri[1]*
https://orcid.org/0000-0002-7554-6505

[1]Banaras Hindu University, MMV, Department of Botany, Varanasi, Uttar Pradesh, India; [2]Banaras Hindu University, Institute of Science, Department of Botany, Varanasi, Uttar Pradesh, India

Received: 2017.11.22; Accepted: 2019.05.19.

*Correspondence: neelammmv14@gmail.com; Tel.: +91- 542-2575246

HIGHLIGHTS

- 144 putative SOD homologs were identified among 85 sequenced cyanobacterial genomes.
- Gene gain-and-loss is insignificant during SOD evolution as they lack additional domain.
- Increased transcript level under abiotic stress confirms their role in abiotic stress.

Abstract: Cyanobacteria are photoautotrophic prokaryotes capable to grow in diverse ecological habitats, originated 2.5-3.5 billion years ago and were first to produce oxygen. Since then superoxide dismutases (SOD) acquired great significance due to their ability to catalyze detoxification of byproducts of oxygenic photosynthesis i.e. superoxide radicals. In the present study, we extracted information regarding SODs from species of sequenced cyanobacteria and investigated their diversity, conservation, domain structure, and evolution. 144 putative SOD homologs were identified. Unlike other protein families (ex.

serine-threonine kinases) SODs are present in all cyanobacterial species reflecting their significant role in survival. However, their distribution varies fewer (0.01%-0.09%) found in unicellular marine strains whereas abundant (0.02%-0.07%) in filamentous nitrogen-fixing cyanobacteria. They were classified into three major subfamilies according to their domain structures: Fe/MnSOD, Cu/ZnSOD and NiSOD. Interestingly, they lack additional domains as found in proteins of other families however motifs and invariant amino acids typical in eukaryotic SODs were conserved well in these proteins indicating similar catalytic mechanism as eukaryotic SODs. Phylogenetic relationships correspond well with phylogenies based on 16S rRNA and clustering occurs on the basis of structural characteristics such as domain organization. Gene gain-and-loss is insignificant during SOD evolution as evidenced by the absence of additional domain. This study has not only examined an overall background of sequence-structure-function interactions for the SOD gene family but also revealed variation among SOD distribution based on ecophysiological and morphological characters.

Keywords: Superoxide dismutases, Cyanobacteria, Comparative genomics, Phylogeny.

INTRODUCTION

Superoxide dismutases (SODs) constitute the first line of defense against oxidative stress in living organisms [1]. SODs constitute a superfamily of metalloenzymes that play a pivotal role in dismutation of highly reactive superoxide radicals thus forestalling generation of various other deleterious reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2), hydroxyl radical ($^{\cdot}OH$), and hypochlorite (OCl^-) [1]. Various abiotic and biotic perturbations (drought, salinity, heavy metal, UV-B, extreme of temperatures, diseases and pests) often cause increasesd generation of ROS in cells [2]. ROS are well known for their damaging effects on membrane, DNA damage, protein oxidation and even lead to severe metabolic dysfunction [2,3]. Therefore to combat ROS toxicity, organisms have developed highly efficient and complex antioxidative defense systems composed of various enzymatic and non-enzymatic components. SODs hold the major position among enzymatic components and catalyze the dismutation of superoxide radicals thereby protecting cells from oxidative damage. Cyanobacteria, a group of photosynthetic oxygen evolving prokaryotes, inhabiting diverse habitats originated around 2.5-3.5 billion years ago [4,5]. They mark presence all over the world and are the only organism liable for making our planet oxidative. Apart from oxygen contribution, they play crucial role in 1) carbon dioxide sequestration, 2) nitrogen fixation 3) and primary productivity in terms of biomass. The group encompasses a large number of species harboring varied genome sizes (1.6-9.0 Mb) of which complete genome sequence of eighty five are available till date. With the increase in number of sequenced cyanobacterial genomes as a result of novel sequencing techniques, new opportunities are approaching for comparative genome research.

Furthermore, cyanobacteria display considerable morphological and ecological diversity. Cell organization pattern is diverse and ranges from unicellular to differentiated multicellular forms. Similarly they are present in diverse habitats such as marine, freshwater and terrestrial environment ranging from polar to tropical climate zones. Among unicellular forms, *Synechocystis* (freshwater), *Synechococcus* and *Prochlorococcus* (marine) are major primary producers of aquatic ecosystem with genome sizes ranging between 1.6 Mb to 3.5 Mb. Various other unicellular species include *Thermosynechcoccus elongatus* BP-1 (2.5 Mb), *Halothece* and *Microcystis* inhabiting hot spring, hypersaline and freshwater ecosystems respectively. However few unicellular genera have comparatively larger genomes, for instance *Cyanothece* sps. (4.7 Mb to 7.8 Mb), *Crocosphaera* (6.3 Mb) etc.

The diazotrophic filamentous forms possess the largest genomes among all cyanobacteria, includes nitrogen fixing fresh water forms such as *Anabaena variabilis* ATCC29413, *Nostoc* PCC 7120 etc and marine forms such as *Trichodesmium erythraeum* IMS101 inhabiting tropical and subtropical oceans. Few genera form symbiotic

relationships with plants for example, *Nostoc punctiforme* ATCC29133. The strain was isolated from symbiotic association with the gymnosperm cycad *Macrozamia* sp.[6].

On the basis of metal cofactor binding there are four isoforms of SODs viz. FeSOD, MnSOD, NiSOD and Cu/ZnSOD. Many SODs from cyanobacteria have been biochemically characterized for instance; FeSOD from filamentous model cyanobacteria *Anabaena* PCC 7120 was reported to be a cytosolic, homodimeric and acidic enzyme exhibiting the characteristic iron peak at 350 nm in its ferric state, an almost 100% occupancy of iron per subunit [7]. Expression analysis of SODs from *Synechocystis* sp. strain PCC 6803 and *Anabaena* have been also carried out [8]. Now, with the availability of genome sequences genome-wide identification have become possible for gene families. As per old version of cyanobase with availability of thirty eight genome sequences, genome-wide identification of serine threonine protein kinases [9], peroxiredoxins [10], carotenoid cleavage dioxygenases [11] and metacaspases [12] family have been carried out. Comparative genomic investigations of cyanobacterial SODs have also been conducted focusing on its structural aspects[13]. However studies targeting comparative analysis based on genome size variation, phylogeny and evolution is lacking. In present study ten previously characterized SODs were selected from *Arabidopsis thaliana*, *Synechococcus* PCC 7942 and *Anabaena* PCC 7120 for blast search at genome level focusing on their classification, distribution, phylogeny and evolution. A better understanding of crucial players (SODs) of antioxidant defense system will help us in unveiling the underlying mechanism.

MATERIALS AND METHODS

Maintenance of cyanobacterial strains

Cyanobacterial strains (*Anabaena* PCC 7120 and *Synechococcus elongatus* 7942) were grown photoautotrophically in BG-11medium buffered with 10 mM HEPES-NaOH, pH 7.5 at 24±2°C under day light fluorescent tubes emitting 72 µmol photon m^{-2} s^{-1} PAR (photosynthetically active radiation) light intensity with a photoperiod of 14:10 h.

Identification of *sod* genes encoding SOD proteins

Eighty five species of cyanobacteria, including Acaryochloris, Calothrix, Chlorobium, Prochlorococcus, Synechococcus, Synechocystis, Gloeobacter, Gloeocapsa, Halothece, Cyanothece, Microcystis, Trichodesmium, Anabaena, Oscillatoria and Nostoc were used in this study. Above mentioned 85 cyanobacterial genomes were downloaded from Cyanobase (new version) [14]. Seven photosynthetic eukaryotic SODs from Arabidopsis thaliana was also downloaded from NCBI Genbank[15]. Moreover, to construct a query protein set known cyanobacterial superoxide dismutases from Synechococcus elongatus 7942 and Anabaena PCC 7120 (Synpcc7942_0801, all0070 and alr2938) were also used. Thus the query set of ten proteins included photosynthetic eukaryotic SODs from Arabidopsis thaliana, Synpcc7942_0801 from Synechococcus elongatus 7942 and all0070, alr2938 from Nostoc PCC 7120 (see supplementary file 1). All SOD genes were searched locally through conducting BLASTp [16-18] and tBLASTn [19] programs from all 85 cyanobacterial genomes using a threshold e-value of 1e- 10. Subsequently, we manually checked the extracted proteins by NCBI CDD, SMART and Pfam analyses to avoid false positives that usually arise during large-scale analyses. SODs found during this analysis were added to the query set for one more round of BLASTp searches. This procedure was repeated till no new proteins were found. All translated protein sequences of genes encoding SODs used in this paper were listed in more detail (see supplementary file 2).

Multiple sequence alignment and structure analysis

Multiple sequence alignment of proteins identified by BLAST was done using ClustalW [20,21] with a gap opening penalty of 10, a gap extension penalty of 0.2, and Gonnet as the weight matrix. Moreover, the SMART [22] and CDD [23] databases were applied to

delete false positives. Furthermore, the alignment was then inspected by analysis of the Fe/MnSOD, Cu/ZnSOD and NiSOD domains [CDD: cl27368, cl00891, cl07609] in the NCBI Conserved Domain Database [23]. A protein was recognized as SOD if any domain mentioned above was identified. Structural analysis of the obtained SODs was performed using the SMART (Simple Modular Architecture Research Tool) [22] and the CDD (Conserved Domains Database) [23], methods, relying on hidden Markov models and Reverse Position- Specific BLAST, respectively. SMART (a Simple Modular Architecture Research Tool) allows the identification and annotation of genetically mobile domains and the analysis of domain architectures based on submitted protein sequence. CD-Search is NCBI's interface to searching the conserved domain database with protein sequences. It uses RPS-BLAST, a variant of PSI-BLAST, to quickly scan a set of pre-calculated position-specific scoring matrices (PSSMs) with a protein query.

Phylogenetic analysis

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6 using maximum likelihood method [24]. This method allows the testing of hypotheses about the constancy of evolutionary rates by likelihood ratio tests, and gives rough indication of the error of the estimate of the tree. Bootstrap probabilities were estimated with 1,000 replications using complete deletion for gaps and missing data. Graphical representation and edition of the phylogenetic tree were also performed with MEGA 6.0.

Real-time quantitative RT-PCR analysis

Total RNA extraction was performed from *Synechococcus elongatus* 7942 and *Anabaena* sp. PCC 7120 before and after 1 day of NaCl (600mM for *Synechococcus elongatus* 7942 and 100mM for *Anabaena* sp. PCC 7120) [25,26] and methyl viologen (50µM for *Synechococcus elongatus* 7942 and 2µM for *Anabaena* sp. PCC 7120) [25,27] treatment using the TRIzol reagent (Invitrogen Inc., CA, USA). cDNA was synthesized from the RNA by using a iScript cDNA synthesis kit (BioRad) according to the manufacturer's instructions. Gene specific primers were designed using primer3 software [28] (supplementary table S1). Reactions were performed in triplicate in a total volume of 20µl including 10 pmol of forward and reverse primers and 1x Sso fast evagreen qPCR supermix (BioRad). A housekeeping gene (*16s*) was used as a reference for normalization and analysed in CFX-96 (Bio-Rad). The comparative ΔΔCt method was used to evaluate the relative quantities of each amplified product in the samples.

RESULTS

Identification of superoxide dismutases

The 85 sequenced cyanobacterial genomes available from Cyanobase were used for this analysis. Phylogeny of eighty five cyanobacterial strains is shown in Figure 1. A total of 144 protein sequences from 85 cyanobacterial genomes are accepted as superoxide dismutases after BLASTP. CDD and SMART analysis were performed to eliminate false positives. Supplementary table S2 displays 144 proteins in detail, among them 128 were annotated as superoxide dismutase and remaining 16 are annotated as putative superoxide dismutase. Most of the proteins lack any additional domains. Only six proteins, ANA C10606, AA 65012270, Cal 7507_0532, Mic 7113_3792, Syn 7502_00221 and tll 1519 contains additional domain of 'phage portal protein' superfamily. One protein Osc 7112_0632 contains an additional ubiquitinol-cytochrome C reductase Fe-S subunit TAT signal.

Figure 1. Phylogenetic tree of the sequenced cyanobacterial strains and SOD information. A phylogenetic tree for 85 sequenced cyanobacteria constructed based on 16s rRNA as was described in methods. Numbers appearing at the nodes corresponded to the values produced by bootstrap analysis (1000 replicates).

The number of SOD genes varies substantially from one to four that is far less than other family of proteins (Table 1). For example number of serine threonine kinases ranges from 0 to 56 in cyanobacterial genomes [9]. Only four cyanobacterial strains Acaryochloris marina MBIC11017, Chroococcidiopsis thermalis PCC 7203, Gloeobacter violaceous PCC 7421 and Stanieria cyanosphaera PCC 7437 harbors four SOD genes. Percentage of SODs ranges 0.02 to 0.07 for filamentous cyanobacteria, however it ranges between 0.01 to 0.09 for unicellular cyanobacteria (supplementary figure S1).

Table 1. Distribution of superoxide dismutases in different cyanobacterial species

Name of species	Total no. of genes	Genome size (Mb)	Total SOD	%
Acaryochloris marina MBIC 11017	8462	8.36	4	0.04
Anabaena cylindrica PCC 7122	5914	7.06	2	0.03
Anabaena sp.90	4570	5.30	1	0.02
Anabaena sp. wa 102	4801	5.78	1	0.02
Anabaena variabilis ATCC 29413	5768	0.74	2	0.03
Calothrix sp.336/3	5108	6.42	2	0.03

Calothrix sp.PCC 6303	5591	6.96	2	0.03
Calothrix sp.PCC 7507	6034	7.02	2	0.03
Candidatus thalassa isolate ALOHA	1242	1.44	1	0.08
Chlorobium tepidum TLS	2255	2.15	1	0.04
Chroococcidiopsis thermalis PCC 7203	5810	6.68	4	0.06
Crinalium epipsammum PCC 9333	4809	5.62	2	0.04
Cyanobacterium aponinum PCC 10605	3486	4.17	2	0.05
cyanobacterium endosymbiont of Epithemia turgida isolate EtSB Lake Yunoko	1993	2.79	1	0.05
Cyanobacterium stanieri PCC 7202	2892	3.16	1	0.03
Cyanobium gracile PCC 6307	3334	3.34	2	0.05
Cyanothece ATCC 51142	5359	5.46	1	0.01
Cyanothece PCC 7424	5767	6.55	1	0.01
Cyanothece PCC 7425	5384	5.78	2	0.03
Cyanothece PCC 7822	6702	7.84	2	0.02
Cyanothece PCC 8801	4420	4.78	2	0.04
Cyanothece PCC 8802	4496	4.80	2	0.04
Dactylococcopsis salina PCC 8305	3427	3.78	2	0.05
Geitlerinema sp.PCC 7407	3873	4.68	2	0.05
Gloeobacter kilaueensis JS1	4562	4.72	3	0.06
Gloeobacter violaceous PCC 7421	4431	4.65	4	0.09
Gloeocapsa sp.PCC 7428	5061	5.88	2	0.03
Halothece sp. PCC 7418	3766	4.17	2	0.05
Leptolyngbya sp.PCC 7376	4281	5.12	2	0.04
Microcoleus sp. PCC 7113	6529	7.96	2	0.03
Microcystis aeruginosa NIES 2549	4329	4.29	2	0.04
Microcystis aeruginosa NIES 843	6363	5.84	2	0.03
Microcystis pannifomis FACHB-1757	6022	5.68	2	0.03
Nostoc azollae 0708	3710	5.48	2	0.05
Nostoc sp. PCC 7107	5329	6.32	3	0.05
Nostoc sp. PCC 7120	6135	7.21	2	0.03
Nostoc sp. PCC 7524	5533	6.71	2	0.03
Nostoc punctiforme PCC 73102	6,794	9.05	3	0.04
Oscillatoria acuminata PCC 6304	5892	7.80	2	0.03
Oscillatoria nigroviridis PCC 7112	6441	8.27	2	0.03
Pleurocapsa sp.PCC 7327	4324	4.98	2	0.04
Prochlorococcus marinus str AS9601	1964	1.66	1	0.05
Prochlorococcus marinus str MIT 9215	2025	1.73	1	0.04
Prochlorococcus marinus str MIT 9301	1949	1.64	1	0.05
Prochlorococcus marinus str MIT 9303	3049	2.68	1	0.03
Prochlorococcus marinus str MIT 9312	2007	1.70	1	0.04
Prochlorococcus marinus str MIT 9313	2966	2.41	1	0.03
Prochlorococcus marinus str MIT 9515	1948	1.70	1	0.05
Prochlorococcus marinus str MIT NATL1A	2236	1.86	1	0.04
Prochlorococcus marinus str MIT NATL2A	2207	1.84	1	0.04
Prochlorococcus subsp *marinus* str CCMP1375	1890	1.75	1	0.05
Prochlorococcus marinus pastoris CCMP 1986	2042	1.65	1	0.04
Prochlorococcus SP MIT 0604	2102	1.78	1	0.04
Prochlorococcus SP MIT 0801	2330	1.92	1	0.04
Pseudoanabaena sp PCC 7367	3909	4.88	3	0.07
Rivularia sp PCC 7116	6710	8.72	3	0.04
Stanieria cyanosphaera PCC 7437	4833	5.54	4	0.08
Synechococcus elongatus PCC 6301	2580	2.69	1	0.03
Synechococcus elongatus PCC 7942	2714	2.74	1	0.03
Synechococcus sp. CC9311	2944	2.60	2	0.06
Synechococcus sp. CC9605	2692	2.51	2	0.07
Synechococcus sp. CC9902	2355	2.23	2	0.08
Synechococcus sp. JA-2-3B''a(2-13)	2919	3.04	1	0.03
Synechococcus sp. JA-3-3Ab	2820	2.93	1	0.03
Synechococcus sp. KORDI-100	3105	2.78	1	0.03
Synechococcus sp. KORDI-49	2783	2.58	1	0.03
Synechococcus sp. KORDI-52	2875	2.57	2	0.06
Synechococcus sp. PCC 6312	3593	3.72	2	0.05

Synechococcus sp. PCC 7002	3235	3.40	1	0.03
Synechococcus sp. PCC 7502	3369	3.58	1	0.02
Synechococcus sp. RCC307	2583	2.22	2	0.07
Synechococcus sp. UTEX 2973	2969	2.73	1	0.03
Synechococcus sp. WH 7803	2586	2.36	2	0.07
Synechococcus sp. WH 8103	2878	2.42	1	0.03
Synechococcus sp. WH 8109	2697	2.11	2	0.07
Synechocystis sp. PCC 6714	3848	3.73	1	0.02
Synechocystis sp. PCC 6803 substr. GT-I	3219	3.57	1	0.03
Synechocystis sp. PCC 6803 substr. PCC-N	3218	3.57	1	0.03
Synechocystis sp. PCC 6803 substr. PCC-P	3218	3.57	1	0.03
Synechocystis sp. PCC 6803 (GCA_001318385.1)	3284	3.56	1	0.03
Synechocystis sp. PCC 6803 (GCA_000340785.1)	3610	3.94	1	0.02
Synechocystis sp. PCC 6803 (GCA_000270265.1)	3220	3.57	1	0.03
Synechocystis sp. PCC 6803 (GCA_000009725.1)	3661	3.94	1	0.02
Thermosynechcoccus elongatus BP-1	2476	2.59	2	0.08
Trichodesmium erythraeum IMS101	4499	7.75	2	0.04

Among unicellular cyanobacteria Candidatus Atelocyanobacterium thalassa isolate ALOHA (0.08%), Stanieria cyanosphaera PCC 7437 (0.08%), Synechococcus sp. CC9902 (0.08%), Thermosynecoccus elongatus BP-1 (0.08%) and Gloeobacter violaceous PCC 7421 (0.09%) harbored maximum percentage of SODs. The number of SOD genes varies with genome sizes (Figure 2). The number of SODs is increasing along with the increase in genome sizes in general however few exceptions exist. This section may be also divided by subheadings. It should provide a concise and accurate description of the experimental results, their interpretation as well as the experimental conclusion that can be drawn.

Structure and Functions

Based on conserved domain database analysis of all SODs, the identified SODs could be classified into three major subfamilies I) Fe/MnSOD, II) Cu/ZnSOD and III) NiSOD. Percent distribution of all cyanobacterial SODs among all three SOD subfamilies was determined which demonstrated that cyanobacterial SOD subfamily I (Fe/MnSOD) includes ninty five SOD (65.9% of total) (supplementary figure S2). Fe and MnSOD are so similar that they have been grouped in one subfamily and due to their high similarity they are also accepted to be arisen from a common ancestor.

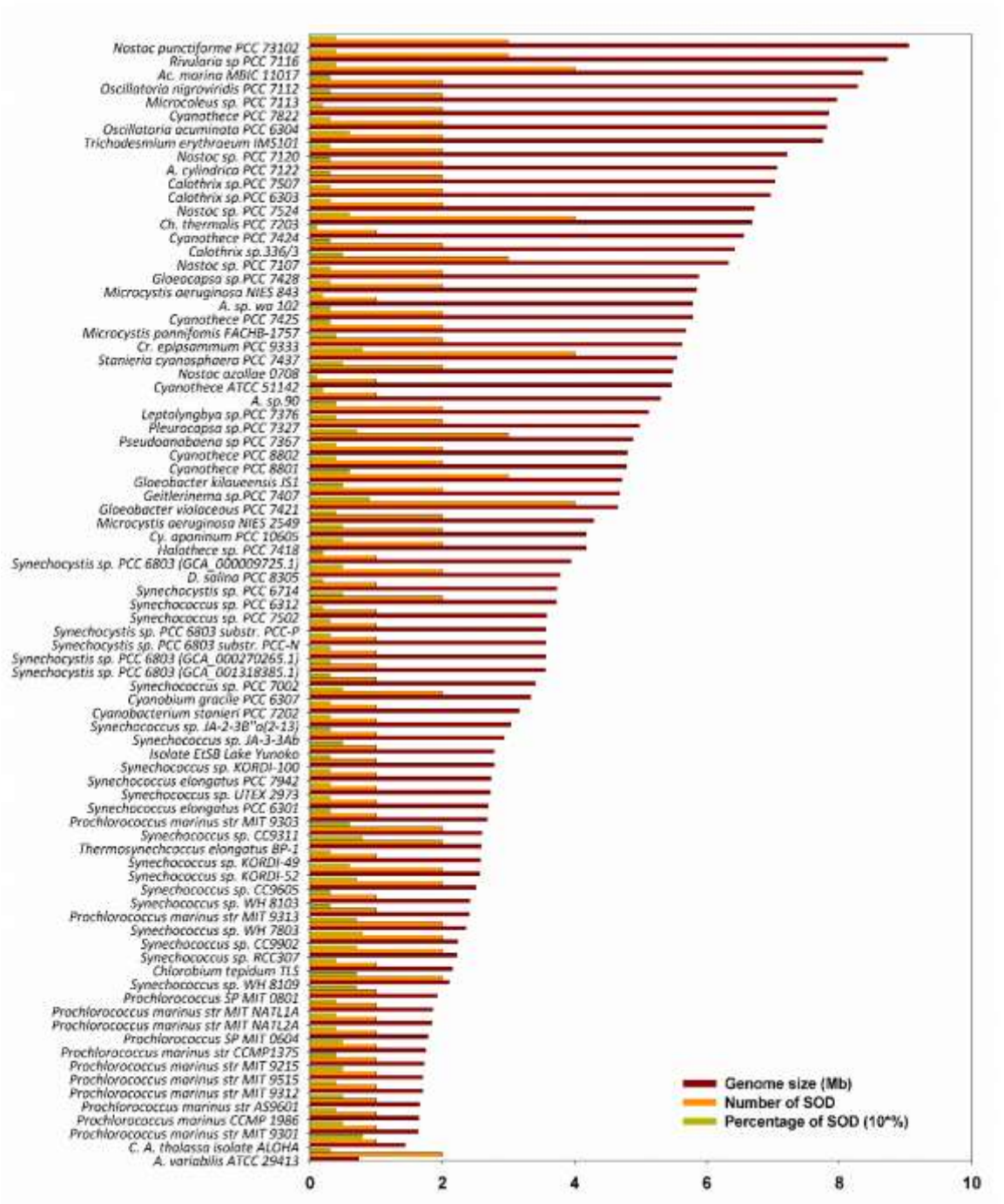

Figure 2. Correlation between the distribution of SOD and the eco-physiological properties and genome sizes of cyanobacteria.

Clustering of Fe and MnSOD in one clade suggests their common origin. Genes encoding SOD from subfamily I are distributed in thirty nine unicellular and two filamentous strains. Subfamily II of SODs (Cu/ZnSOD) includes eighteen members distributed among twelve unicellular (*Acaryochloris marina* MBIC11077, *Chroococcidiopsis thermalis* PCC 7203, *Stanieria cyanosphaera*, *Gloeobacter kilauensis* JS1, *Gloeobacter violaceus* PCC 7421 and seven strains of *Synechococcus* sp.) and four filamentous genera (*Geitlerinema* sp. PCC 7407, *Leptolyngbya* sp. PCC 7316, *Nostoc* sp. PCC 7107, and *Crinalium epipsammum* PCC 9333). Moreover, third subfamily of SODs i.e. NiSOD contains thirty one members distributed among twenty six unicellular and five filamentous genera. Percent distribution of different subfamilies of SOD among unicellular and filamentous strains suggests dominance of NiSOD in unicellular strains however filamentous strains contains large percentage of subfamily I (Fe/MnSOD) (supplementary fig S3).

Phylogenetic Analysis

To explicate the evolutionary relations between cyanobacterial superoxide dismutases, the translated fasta sequences of all genes were subjected to construct the phylogenetic tree (Figure 3).

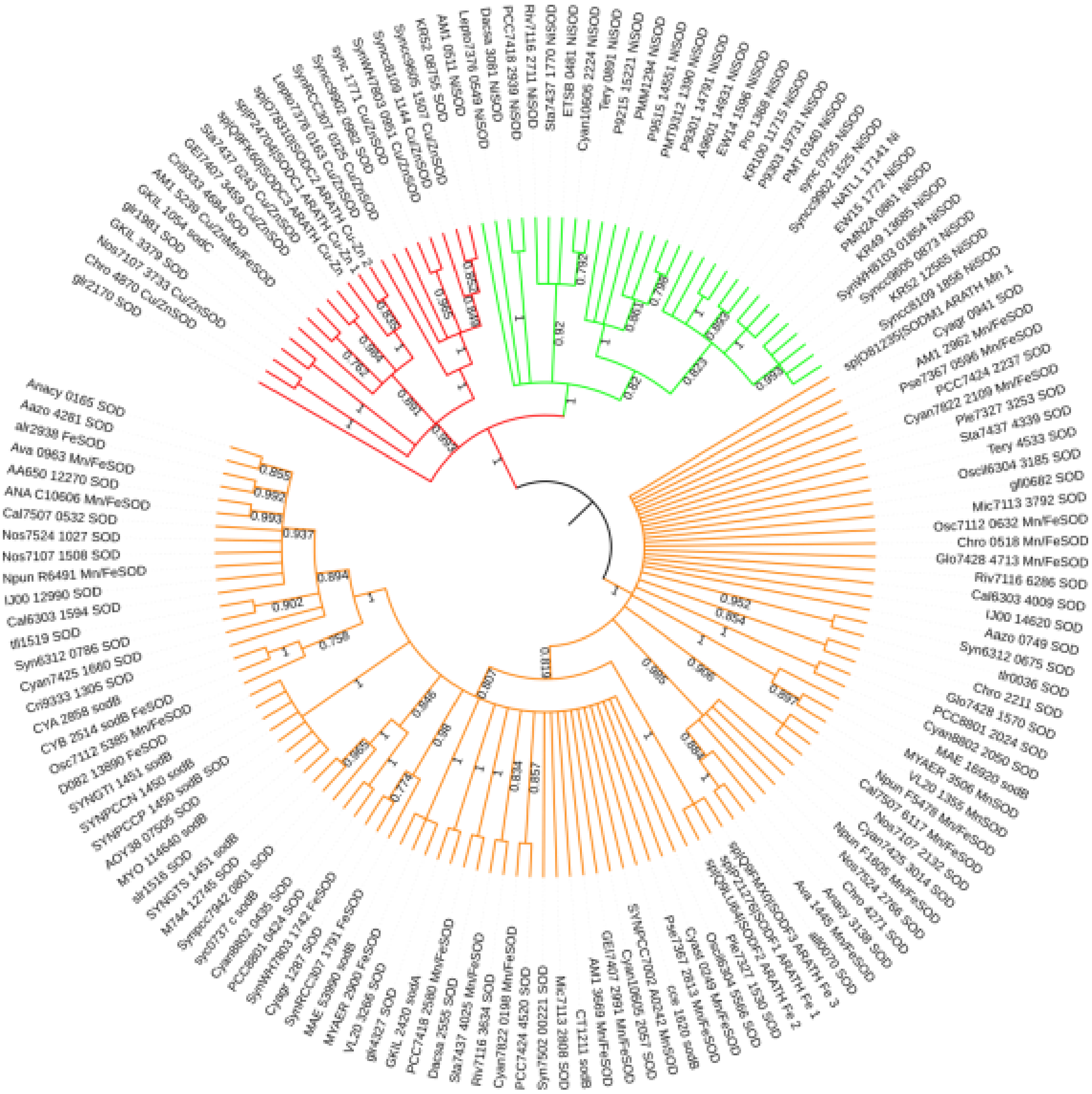

Figure 3. Phylogenetic trees of the total SODs. The evolutionary history was inferred using the Neighbor-Joining method. The evolutionary distances were computed using the p-distance method and are in the units of the number of amino acid differences per site. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 133 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.

In general, three major clades were observed in the phylogenetic tree. It clearly demonstrates grouping of cyanobacterial SODs according to their classification based on CDD analysis, therefore it can be concluded that clustering of cyanobacterial SODs takes place according to their structural characteristics. SODs from the unicellular marine *Prochlorococcus* and *Synechococcus* group together and displays close resemblance with other unicellular strains. Similarly, SODs from filamentous diazotrophic *Anabaena*, *Nostoc* and *Calothrix* clusters together in all three clades of respective SOD subfamilies. This trend is consistent with the clustering pattern followed in 16S rRNA tree. Interestingly, chloroplast SOD F1, F2 and F3 clusters together with Fe/MnSOD from filamentous

cyanobacteria, suggesting a close evolutionary relationship between chloroplastic genes of higher plant and cyanobacteria thus witnessing the cyanobacterial origin of chloroplast.

Conserved domain features

To elucidate the conserved domain features and motifs of the cyanobacterial SODs, multiple sequence alignment was performed.

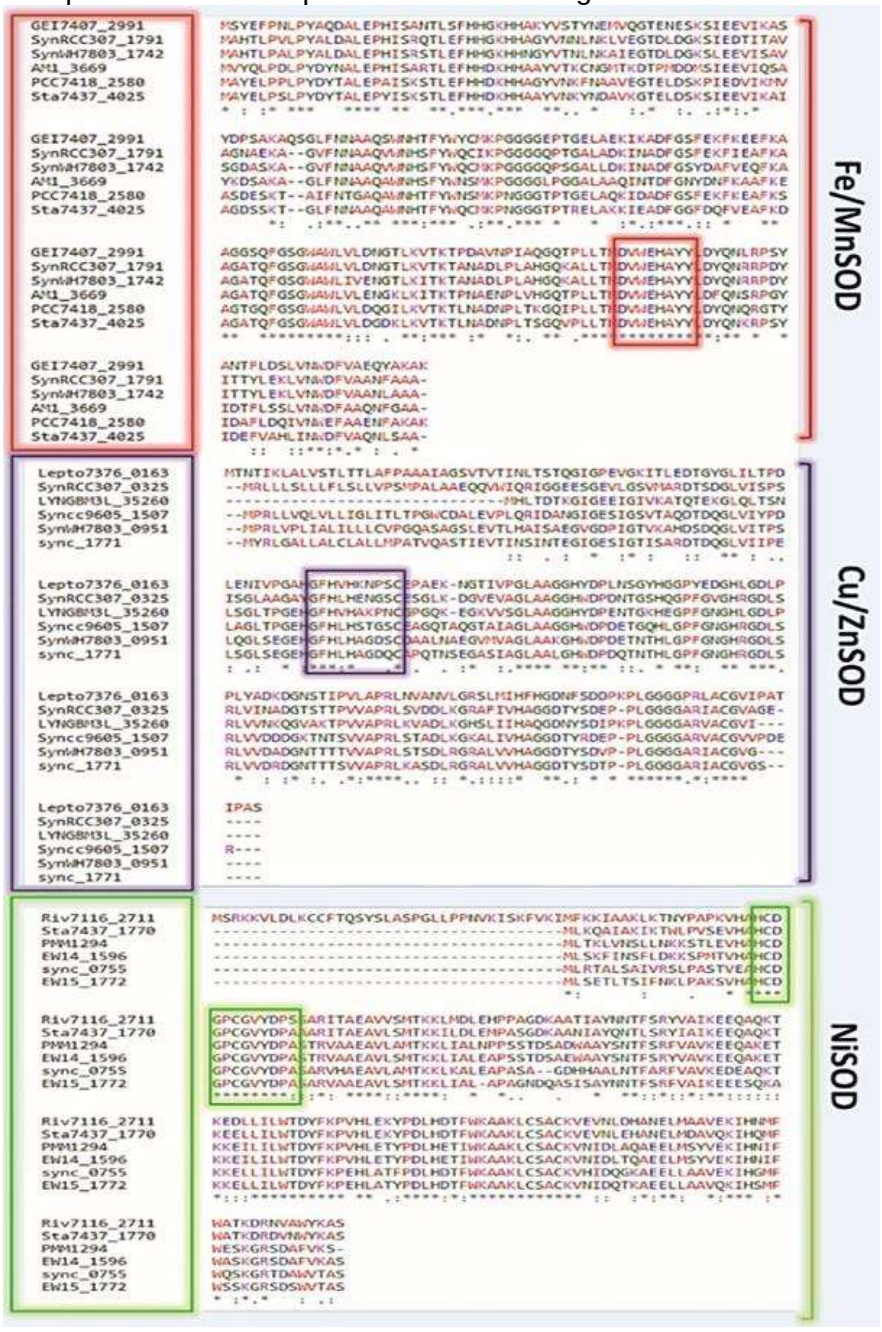

Figure 4. Conserved motifs in cyanobacterial superoxide dismutases from different subfamilies

FeSOD are well known for presence of conserved metal-binding domain “DVWEHAYY” [29] as also reflected in present study. MnSOD and FeSOD shares N and C terminal domains as revealed by CDD analysis. The metal binding motif of Mn/FeSOD is shown in multiple sequence alignment (Figure 4). Similarly multiple sequence alignment revealed signature sequence of Cu/ZnSOD as G-F-H-[ILV]-H-x-[NGT]-[GPDA]-[SQK]-C. Furthermore, NiSOD contained conserved motif “HCDGPCVYDPA”. Interestingly, unlike various other cyanobacterial gene families (peroxiredoxins and serine threonine kinases) additional domains are lacking in cyanobacterial SODs except seven proteins ANA C10606 (*Anabaena* sp. 90), AA 65012270 (*Anabaena* sp. wa102), Cal 7507_0532

(*Calothrix* sp. PCC 7507), Mic 7113_3792 (*Microcoleus* sp. PCC 7113), Syn 7502_00221 (*Synechococcus* sp. PCC 7502), tll 1519 (*Thermosynechococcus elongatus* BP-1) and Osc 7112_0632 (*Oscillatoria nigro-viridis* PCC 7112).

Differential expression pattern of gene in response to abiotic stress

The expression pattern of SODs from two model organism, *Synechococcus elongatus* 7942 and *Anabaena* PCC 7120 respectively representing unicellular and filamentous forms was detected under two abiotic stresses (methyl viologen and salinity). The results showed that transcript of SOD genes (*all0070* and *alr2938*) from *Anabaena* PCC 7120 were upaccumulated 3.5 fold and 8 fold respectively under methyl viologen stress and 10 and 20 fold respectively under salt stress. Furthermore, SOD gene (Synpcc7942_0801) from unicellular cyanobacteria *Synechococcus* displayed 1.6 and 4 fold upaccumulation under methyl viologen and salt stress respectively (Figure 5).

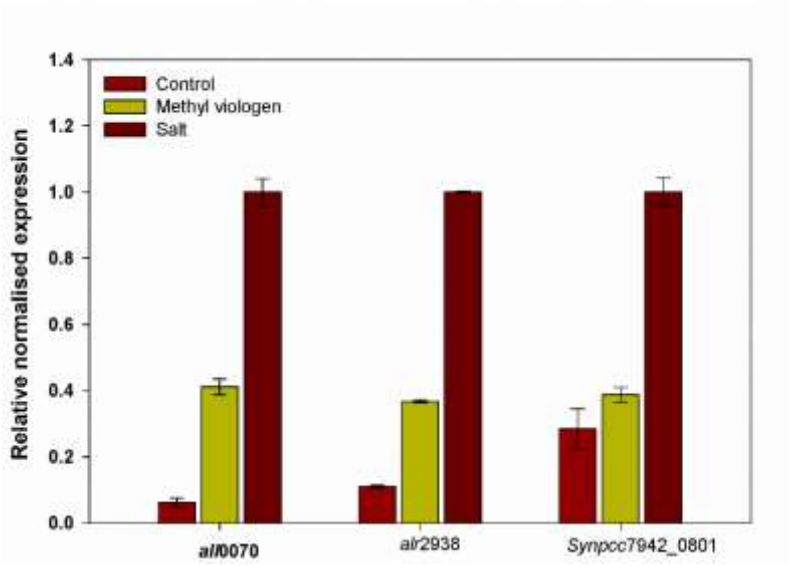

Figure 5. Relative normalized expression of superoxide dismutases from *Anabaena* PCC 7120 and *Synechococcus elongatus* 7942 under methyl viologen and salt stress.

DISCUSSION

Superoxide dismutases have acquired great significance following emergence of oxygenic photosynthesis due to their intrinsic ability to catalyze the detoxification of superoxide radicals. Therefore it is very vital to study them in cyanobacteria which originated 2.5-3.5 billion years ago and brought oxygenic photosynthesis, that lead to transition of environment (reducing to oxidising). Their essential role in mitigating oxidative stress is successfully demonstrated by many research groups, for instance Thomas et al., 1998 demonstrated sensitivity of *Synechococcus* sp. strain PCC 7942 lacking functional FeSOD to methyl viologen or norfluarazon [30].

The SODs used in present study were identified by BLAST and were manually checked for false negative and positives, a common error arise during large scale automated analysis. CDD, SMART analysis and sequence alignment results displayed cyanobacterial SODs also possess similar conserved signature sequence as eukaryotic SODs. Presence of similar conserved sequences in both cyanobacteria and eukaryotes suggests a quite similar mechanism of action, however variation in domain organization can be seen.

The distribution of SODs is related to the genome sizes and ecological conditions. The distribution of putative SODs encoding open reading frame among various cyanobacteria correlated with their genome sizes however few variation also exists. The exceptions lack the correlation between number of SODs and genome size suggesting that larger genome sizes are not duplication events but are due to acquisition of additional functions. Moreover

another fact from the present study i.e. distribution of small number of SODs in marine cyanobacteria from the present study correlated well with previous studies on serine threonine kinases, peroxiredoxins and metacapsases. In marine cyanobacterial sps. reduction in STKs, Prxs and metacapsases is reported [9,10,12]. The possible reason behind this phenomenon is reported to be a selective force that favors the survival of these cyanobacteria under infavorable marine environment.

Novel proteins are known to be produced either through insertion or shuffling of domains along with gene gain-and-loss, however in case of cyanobacterial SODs only two additional domains viz. phage portal protein and ubiquitinol-cytochrome c reductase Fe-S subunit TAT signal in 7 SODs out of total 144 SODs are present.

Genome-wide identification of SODs in different cyanobacterial genera demonstrates presence of NiSOD in primitive unicellular and less evolved genera supporting the earlier studies. Till now no evidences are found about presence of NiSOD in G^+ bacteria, archaea or eukaryotes thus restricted to relatively few groups and assumed that NiSOD have been evolved after differentiation of eukaryotes. Cu/ZnSOD are very rare in cyanobacterial genomes as reflected by present study and Fe/MnSOD are the most abundant one present in middle order and most evolved forms.

The phylogenetic tree for SODs revealed that cyanobacteria and higher plants share a common ancestor, consistent with earlier studies [9,10]. The phylogenetic relationship among SODs from higher plants and cyanobacteria strongly supports cyanobacterial origin of these proteins in higher plants, indicating possible gene acquisition from cyanobacteria by endosymbiosis event. Furthermore, phylogenetic tree based on amino acid sequences of SODs coincide well with the phylogenies based on the 16S rRNA. All 3 types of cyanobacterial SODs (Fe/MnSOD, Cu/ZnSOD and NiSOD) are reported to be involved in ROS scavenging caused by abiotic stress. A comparative expression from model organism of filamentous form (*Anabaena* PCC 7120) and unicellular form (*Synechococcus elongatus* 7942) under abiotic stress was performed. *Anabaena* PCC 7120 harbours 2 SOD genes (all0070 and alr2938) whereas *Synechococcus* genome contains only one SOD (Synpcc7942_0801). A comparative expression analysis of all 3 genes under methyl viologen and salt stress displays maximum upregulation of alr2893, followed by all0070 and Synpcc7942_0801. This indicates that alr2938 might play a predominant antioxidant role in *Anabaena* PCC 7120.

Funding: This research received no external funding.
Acknowledgments: Shivam Yadav and Rajesh Prajapati thanks to the University Grants Commission, New Delhi for SRF and RGNF respectively. We are also thankful to the Head of the department and Coordinator CAS in Botany for facilities.
Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

1. McCord J, Fridovich I. Superoxide dismutase: an enzymic function for erythrocuprein (hemocuprein). *J Biol Chem.* 1969; 244: 6049–6055.
2. Karuppanapandian T, Moon J-C, Kim C, Manoharan K, Kim W. Reactive oxygen species in plants: their generation, signal transduction, and scavenging mechanisms. *Aust J Crop Sci.* 2011; 5: 709–25.
3. Lee SH, Ahsan N, Lee KW, Kim DH, Lee DG, Kwak SS. Simultaneous overexpression of both CuZn superoxide dismutase and ascorbate peroxidase in transgenic tall fescue plants confers increased tolerance to a wide range of abiotic stresses. *J Plant Physiol.* 2007; 164: 1626–38.

4. Blankenship RE, Hartman H. The origin and evolution of oxygenic photosynthesis. *Trends Biochem Sci*. 1998; 23: 94–97.
5. Whitton BA, Potts M. The ecology of cyanobacteria: their diversity in time and space. Whitton, B.A.; Potts, M. 2000; pp. 1–11.
6. Rippka R, Herdman M. Pasteur culture collection of cyanobacterial strains in axenic culture, Catalogue and taxonomic handbook. Institut Pasteur, Paris. 1992; pp.103
7. Regelsberger G, Laaha U, Dietmann D, Ruker F, Canini A, Grilli-Caiola M, et al. The iron superoxide dismutase from the filamentous cyanobacterium *Nostoc* PCC 7120: localization, overexpression, and biochemical characterization. *J Biol Chem.* 2004; 279: 44384 –44393.
8. Kim JH, Suh KH. Light-dependent expression of superoxide dismutase from cyanobacterium *Synechocystis* sp. strain PCC 6803. *Arch Microbiol.* 2005; 183: 218–223.
9. Zhang XW, Zhao FQ, Guan XG, Yang Y, Liang CW, Qin S. Genome-wide survey of putative serine/threonine protein kinases in cyanobacteria. *BMC Genomics.* 2007; 8: 395
10. Cui H, Wang Y, Wang Y, Qin S. Genome-wide analysis of putative peroxiredoxin in unicellular and filamentous cyanobacteria. *BMC Evol Biol.* 2012;12: 220
11. Cui H, Wang Y, Qin S. Genome-wide analysis of carotenoid cleavage dioxygenases in unicellular and filamentous cyanobacteria. *Comp Funct Genomics.* 2012; 1-13.
12. Jiang Q, Qin S, Wu QY. Genome-wide comparative analysis of metacaspases in unicellular and filamentous cyanobacteria. *BMC Genomics* 2010; 11: 198.
13. Priya B, Premanandh J, Dhanalakshmi RT, Seethalakshmi T, Uma L, Prabaharan D, et al. Comparative analysis of cyanobacterial superoxide dismutases to discriminate canonical forms. *BMC Genomics*. 2007; 8: 435
14. Cyanobase: http://genome.kazusa.or.jp/cyanobase.
15. NCBI: http://www.ncbi.nlm.nih.gov/sites/gquery.
16. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* . 1990; 215: 403–410.
17. Mount DW. Using the Basic Local Alignment Search Tool (BLAST). *CSH Protoc.* 2007; pdb top17.
18. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25: 3389–3402.
19. Gertz EM, Yu YK, Agarwala R, Schaffer AA, Altschul SF. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biol*, 2006; 4: 41.
20. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007; 23: 2947–2948.
21. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994; 22: 4673–4680.
22. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci USA*. 1998; 95: 5857–5864.
23. Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 2004; 32: W327–W331.
24. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biol and Evol*. 2013; 30: 2725-2729.

25. Stork T, Laxa M, Dietz MS, Dietz KJ. Functional characterisation of the peroxiredoxin gene family members of *Synechococcus elongatus* PCC 7942. *Arch Microbiol*. 2009; 191: 141-51.
26. Rai S, Agrawal C, Shrivastava AK, Singh PK, Rai LC. Comparative proteomics unveils cross species variations in *Anabaena* under salt stress. *J Proteomics*. 2014; 98: 254–270
27. Panda B, Basu B, Rajaram H, Apte SK. Methyl viologen responsive proteome dynamics of *Anabaena* Sp. Strain PCC 7120. *Proteomics*. 2014; 14: 1895-1904.
28. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol.* 2000; 132: 365–386
29. Feng X, Lai Z, Lin Y, Lai G, Lian C. Genome-wide identification and characterization of the superoxide dismutase gene family in *Musa acuminate* cv. Tianbaojiao (AAA group). *BMC Genomics.* 2015; 16: 823
30. Thomas DJ, Avenson TJ, Thomas JB, Herbert SK. A cyanobacterium lacking iron superoxide dismutase is sensitized to oxidative stress induced with methyl viologen but is not sensitized to oxidative stress induced with norflurazon. *Plant Physiol.* 1998; 116: 1593-1602