

Genome-wide distribution of bilin-binding GAF domains

Shinobu Okamoto¹

okamoto@kuicr.kyoto-u.ac.jp

Masayuki Ohmori²

ohmori@molbiol.saitama-u.ac.jp

Minoru Kanehisa¹

kanehisa@kuicr.kyoto-u.ac.jp

- 1 Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan
- 2 Department of Molecular Biology, Faculty of Science, Saitama University, 255 Shimo-Ohkubo, Sakura-ku, Saitama 338-8570, Japan

Keywords: genome, GAF domain, phytochrome, cyanobacteria

1 Introduction

GAF domains were first described for vertebrate cGMP-specific phosphodiesterases, a cyanobacterial adenylate cyclase and the bacterial formate hydrogen lyase transcription activator FhlA [1]. GAF domains consist of about 150 amino acids conserved in many signaling molecules. GAF domains (PF01590) can be identified in the Pfam database by their characteristic protein sequence signature. GAF domains of plant phytochromes represent the bilin-lyase-specific GAF subfamily [2]. The phytochromes form a major class of signal transducing photoreceptors in plants and the largest group of GAF containing proteins. Although, they were first discovered in higher plants, they are also widely distributed in bacteria.

In this study, we surveyed the distribution of GAF domains in many organisms according to the hypothesis that function as a light receptor.

2 Methods

The genome sequences of the 150 organisms were obtained from the KEGG/GENES, JGI, and GenBank databases. Domain search was performed with the programs HMMER package and the Pfam database (release 7.2) [3]. Cluster analysis was carried out by the ClustalW program.

3 Results and Discussion

The survey of genome sequences of 150 organisms revealed the presence of 1252 GAF domains encoded by 1065 ORFs (Fig. 1A). The 1252 GAFs domains could be categorized into more than ten subclasses based on the cluster analysis. There are 99 ORFs encoding 141 GAFs categorized into phytochrome related CB-GAF domains (Table 1). Surprisingly, 79% of the CB-GAFs exist in cyanobacteria. The 141 CB-GAF domains could be clustered into four classes using the neighbor-joining algorithm (Fig. 1B). Class I is involved in higher plant. Class II is involved in photosynthetic, non-photosynthetic and cyanobacteria. Class III and Class IV consist of cyanobacterial phytochrome-like GAFs.

Based on these results, we propose that the GAF domains were explosively multiplied with functional diversity, and one part of them has evolved into the CB-GAF.

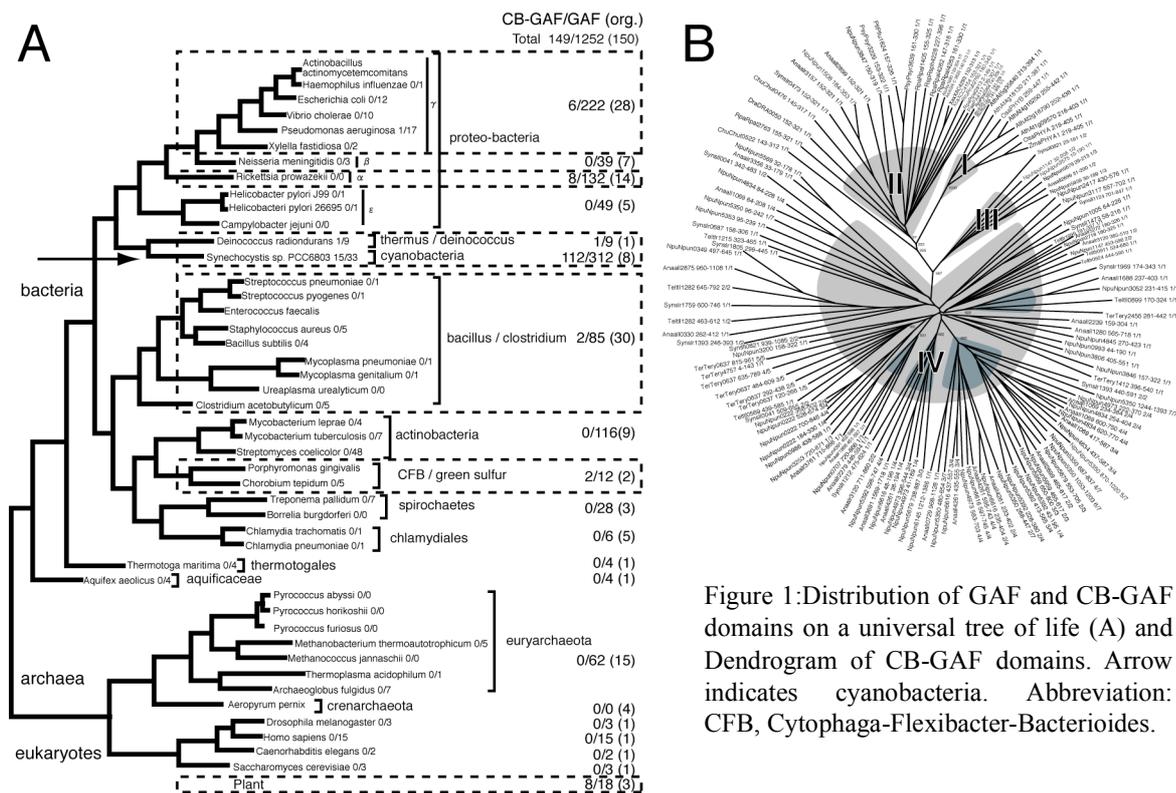


Table 1: Number and distribution of CB-GAF and GAF domains in the genomic database.

Organism	GAF mol ^a	GAF dom ^b	CB mol ^c	CB dom ^d	phylum
<i>Arabidopsis thaliana</i>	10	10	5	5	Plants
<i>Oryza sativa</i> ^e	5	5	2	2	Plants
<i>Zea mays</i> ^e	3	3	1	1	Plants
<i>Synechocystis</i> sp. PCC 6803	28	33	11	14	Cyanobacteria
<i>Anabaena</i> sp. PCC7120	60	85	18	26	Cyanobacteria
<i>Nostoc punctiforme</i>	86	137	30	56	Cyanobacteria
<i>Trichodesmium erythraeum</i> IMS101	19	27	4	8	Cyanobacteria
<i>Thermosynechococcus elongatus</i> BP-1	16	21	7	8	Cyanobacteria
<i>Deinococcus radiodurans</i> R1	8	9	1	1	Deinococcus-Thermus
<i>Agrobacterium tumefaciens</i> C58 (U.Washington/Dupont)	12	12	2	2	Proteobacteria alpha
<i>Agrobacterium tumefaciens</i> C58 (Cereon)	12	12	2	2	Proteobacteria alpha
<i>Rhodospseudomonas palustris</i> CGA009	20	21	6	6	Proteobacteria alpha
<i>Rhodospirillum rubrum</i> ATCC 11170	22	23	1	1	Proteobacteria alpha
<i>Rhodobacter sphaeroides</i> CGA009	13	13	1	1	Proteobacteria alpha
<i>Pseudomonas aeruginosa</i> PAO1	17	17	1	1	Proteobacteria gamma
<i>Pseudomonas fluorescens</i> PFO-1	13	13	1	1	Proteobacteria gamma
<i>Pseudomonas syringae</i> B728a	18	19	2	2	Proteobacteria gamma
<i>Xanthomonas axonopodis</i> pv. Citri 306	20	21	1	1	Proteobacteria gamma
<i>Xanthomonas campestris</i> pv. <i>Campestris</i> ATCC 33913	18	19	1	1	Proteobacteria gamma
<i>Cytophaga hutchinsonii</i> ATCC 33406	7	7	2	2	Cytophaga-Flavobacterium

a) Genes containing GAF domain. b) GAF domain. c) Genes containing CB-GAF domain. d) CB-GAF domain. e) draft genome.

Acknowledgment

This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- [1] Aravind, L. and Ponting, C.P., The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem. Sci.*, 22: 458-459, 1997.
- [2] Wu, S.H. and Lagarias, J.C., Defining the bilin lyase domain: lessons from the extended phytochrome superfamily. *Biochemistry*, 39:13487-13495, 2000.
- [3] Sonnhammer, E.L., et al., Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.*, 26:320-322, 1998.