[Frontiers in Bioscience S6, 1-15, January 1, 2014]

Among the three striatin family members, SG2NA was first to arise during evolution

Goutam Kumar Tanti[1], Nandini Singarapu[1,2], Rohini Muthuswami[1], Shyamal K. Goswami[1]

[1]School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India, [2]Department of Molecular Carcinogenesis, University of Texas M. D. Anderson Cancer Center, Smithville, Texas 78957, USA.

TABLE OF CONTENTS

1. ABSTRACT

Striatin, SG2NA, and zinedin constitute a three-member subfamily of WD-40 repeat proteins. They are found only in metazoans and are likely to have scaffolding functions. Apart from WD-40 repeats, they also have a caveolin-binding motif, a coiled-coil structure, and a calmodulin-binding domain. This paper focuses on the analysis of their evolution as a paradigm of understanding the metazoan scaffolds. Each member of the family forms distinct phylogenetic clusters, wherein striatins, SG2NAs, and zinedins have 13, 10, and 9 conserved motifs, respectively. Furthermore, two of those motifs each in striatin and in zinedin and three in SG2NA are exclusive for the respective subfamily. Of those exclusive motifs for SG2NA, two encompass the caveolin-binding and coiled-coiled domains. Collectively, they show the presence of 11 conserved motifs, suggestive of convergence of individual motifs and creation of patterns. A prokaryotic WD-40 repeat motif pCM-I was found only in the corresponding domain of SG2NA but not in other family members. It is thus hypothesized that striatin family members have evolved from bacteria, and SG2NA was the first member to arise.

2. INTRODUCTION

WD40 repeat superfamily of proteins are characterized by the presence of repetitive 44–60 minimally conserved amino acid residues containing a glycine-histidine (GH) pair at 11–24 position from its N-terminus and a tryptophan–aspartic acid (WD) pair at the C-terminus (1-2). WD-40 is one of the most abundant protein domains in eukaryotes. A conservative estimation suggests that in there are 349 genes encoding WD-40 domains in human (3-4). Structural analyses suggest that WD-40 repeat domains form a propeller-like conformation that provides a scaffolding platform for protein–protein interactions (4). Members of this family participate in biological functions as diverse as ubiquitin-mediated protein turnover (5), chromatin organization (6), cell signaling (7), vesicular trafficking (8), cytoskeletal assembly (9), and cell cycle regulation (10). Considering such diverse functions, it is likely that members of this superfamily have originated at different time points in evolution (4, 11).

Striatin, SG2NA, and zinedin constitute a three-member subfamily of WD-40 repeat proteins, characterized by the presence of a caveolin-binding motif, a coiled-coil

structure, and a calmodulin-binding domain arranged in the same order from the N-terminus to the C-terminus (12). Striatin is the prototype member of the family, first characterized as a calmodulin-binding protein from rat brain synaptosome and in abundance in the striatum—hence named striatin (13). SG2NA (also called striatin-3) was initially identified as a **n**uclear auto-**a**ntigen whose expression is augmented during **S** to **G2** phase of the cell cycle (14). Zinedin, the third member of the family was reported later, and the structural relatedness among three was established (15). All three members of the family have caveolin-binding activities and are considered to be functionally related (12, 15-16). Striatin and SG2NA also interact with protein phosphatase 2A (PP2A), attributing them to vesicular trafficking (17). Several other interacting partners of striatin exist, which suggest a wider network in which respective members of the family have both overlapping and distinctive interactomes (18-20).

Members of the striatin family are conserved in metazoan evolution and are absent in plants and prokaryotes (12). We have demonstrated that mouse and chick SG2NAs have multiple variants arising from alternative splicing (21) and SG2NA is regulated by epigenetic modifications mediated by Brg-1 (22). Some of those variants are devoid of the WD-40 repeats, and their expression profiles vary in different tissues and cell lines [(21) and unpublished results)].

Higher metazoans such as human and mouse have three orthologs: *striatin*, *sg2na* and *zinedin*; *Drosophila melanogaster* and *Caenorhabditis elegans* have only one homolog each: *cka* and *cash-1*, respectively (23). Since metazoan evolution involves more complex interactions between various protein domains, we argue that the four signature domains of the striatin subfamily of WD-40 proteins might be a treasured paradigm of understanding the evolution of signal scaffolds. We have thus undertaken a phylogenetic and sequence conservation study analyzing the functional relatedness among the members of the family. We performed BLASTP analysis of databases available at NCBI and retrieved the homologous protein sequences of various organisms. Gblock analysis is in use for a long time to identify the novel conserved motifs in protein family (24-25). These conserved motifs, in turn, can be used to identify proteins from newly sequenced genomes (24). We used similar approach to identify conserved motifs by using Gblock analysis. Then, ScanProsite analysis was used for their annotation. MAFFT was used for phylogenetic analysis. Results were then analyzed to identify the origin and evolution of the striatin family. Our results suggest that during speciation, a common ancestral gene in lower eukaryotes further diversified into three orthologs of the striatin family, of which *sg2na* is the most ancient.

3. RESULTS

3.1. Striatin-like proteins are conserved in eukaryotes

Members of the striatin family have been reported in organisms ranging from fungi to human, but they are absent in plants and prokaryotes (12). They are expressed in abundance in the central nervous system (CNS) and are thought to be involved in CNS development (12, 26). Considering this, we analyzed the conservation of striatin family members among a range of species, starting from those with a diffused nervous system to those with a well-developed CNS (Tables 1–3).

BLASTP analysis using human striatin, SG2NA, and zinedin enabled us to retrieve their counterparts in both metazoan and nonmetazoan organisms. Further analysis revealed that while striatin and SG2NA are present in almost all metazoans, it is not true for zinedin. Although zinedin has been identified in many mammalian species, it has no homolog identified in the platypus *(Ornithorhynchus anatinus*), the earliest mammal to evolve, which has both striatin and SG2NA homologs but not zinedin homologs. Similarly, among aves, chicken (*Gallus gallus domesticus*) has striatin and SG2NA but is devoid of zinedin (see Table 1). In addition unnamed protein sequences homologous at C-terminus hzinedin were identified in zebra finch (*Taeniopygia guttata*). Striatin-like proteins were also predicted in fishes, such as zebrafish (*Danio rerio*) and Tilapia (*Oreochromis niloticus*), and partially homologous sequences were also identified as unnamed proteins in pufferfish (*Tetraodon nigroviridis*), suggesting the possibility that, similar to SG2NA and striatin, zinedin homologs might be widely distributed among the metazoan, but are yet to be identified. Lower metazoan such as *D. melanogaster* has only one homolog, that is, CKA. Similarly, *H. magnipapillata* (hydra), which has diffuse nervous system, possesses only one homolog of the family.

BLASTP analysis also identified a single homolog in *Schizosaccharomyces pombe* but not in *Saccharomyces cerevisiae* (a complete list of organisms from lower eukaryotes is given in Table 2). While certain species of fungi, such as *Aspergillus niger*, have a striatin-like protein, it is absent in others. A striatin-like protein was also found in *Dictyostelium discoideum*. As expected, no specific hits were found in rice and Arabidopsis genome databases, although there were a substantial number of WD-40 repeat sequences. Bacterial and archeal genome databases also did not harbor any striatin family member, although partially conserved WD-40 repeats were found. For example, in *Thermomonospora curvata*, a WD-40 repeat–containing metallophosphoesterase was identified with an e value of $4e^{-27}$. A complete list of prokaryotes analyzed in this report is given in Table 3.

3.2. Striatin, SG2NA, and zinedin form separate phylogenetic clusters

Phylogenetic analysis showed that striatin, SG2NA, and zinedin form three separate clusters (Figure 1). To further analyze the phylogenetic relationship between metazoan and nonmetazoan homologs, unrooted and rooted phylogenetic trees were created by aligning the N-terminal regions of striatin-like proteins. It was found that striatin family members in higher eukaryotes clustered away from those in lower eukaryotes (Figure 2). Furthermore, the higher eukaryotic homologs of only the N-terminal region of striatin also formed three subclusters composed of striatin, SG2NA, and zinedin, respectively.

Table 1. Higher eukaryotic organisms studied

SN	Protein	Organism	Genera	Scientific name	Abbreviation	Accession no.
1	Striatin	Rat	Mammal	*Rattus norvegicus*	RnoSTRN	P70483.1
2	Striatin	Mouse	Mammal	*Mus musculus*	MmuSTRN	AAI50728.1
3	Striatin	Human	Mammal	*Homo sapiens*	HsaSTRN	NP_003153.2
4	Striatin	Frog	Amphibian	*Xenopus laevis*	XlaSTRN	NP_001087621.1
5	Striatin	Chicken	Bird	*Gallus galus*	GgaSimilarSTRN	XP_419519.2
6	Striatin	Zebrafish	Fish	*Danio rerio*	DreSTRN	NP_001074111.1
7	Striatin	Bovine	Mammal	*Bos Taurus*	BtaSTRN	NP_001193033.1
8	Striatin	Gold fish	Fish	*Carassius auratus*	CauSTRN	ABY50451.1
9	Striatin	Pufferfish	Fish	*T. nigroviridis*	TniSimilarSTRN	CAG09536.1
10	Striatin	Orangutan	Mammal	*Pongo abelii*	PabSTRN	XP_002812184.1
11	Striatin	Platypus	Semi-aquatic mammal	*Ornithorhynchus anatinus*	OanSTRN	XP_001508243.1
12	Striatin	Ant	Insect	*N. vitripennis*	NviSimilarSTRN	XP_001601585.1
13	SG2NA	Human	Mammal	*Homo sapiens*	HsaSG2NA	Q13033.3
14	SG2NA	Rat	Mammal	*Rattus norvegius*	RnoSG2NA	P58405.2
15	SG2NA	Mouse	Mammal	*Mus musculus*	MmuSG2NA	Q9ERG2.1
16	SG2NA	Bovine	Mammal	*Bos Taurus*	BtaSG2NA	A5D7H2.1
17	SG2NA	Chicken	Mammal	*Gallus gallus*	GgaSG2NA	XP_421225.2
18	SG2NA	Base tunicate	Spermatozoa	*Ciona intestinalis*	CinSG2NA	XP_002127249.1
19	SG2NA	Dog	Mammal	*Canis familiaris*	CfaSG2NA	XP_537404.2
		Platypus	Semiaquatic mammal	*Ornithorhynchus anatinus*	OanSG2NA	XP_001512815.1
20	SG2NA					
22	SG2NA	Orangutan	Mammal	*Pongo abelii*	PabSG2NA	XP_002824688.1
23	Zinedin	Mouse	Mammal	*Mus musculus*	MmuZinedin	P58404.1
24	Zinedin	Dog	Mammal	*Monodelphis domestica*	MdoZinedin	XP_001373440.1
25	Zinedin	Human	Mammal	*Homo sapiens*	HsaZinedin	NP_001034966.1
26	Zinedin	Rat	Mammal	*Rattus norvegicus*	RnoZinedin	NP_001100950.2
27	Zinedin	Orangutan	Mammal	*Pongo abelii*	PabZinedin	XP_002829495.1
28	Zinedin	Bovine	Mammal	*Bos Taurus*	BtaZinedin	XP_874777.3
29	Zinedin	European rabbit	Mammal	*Oryctolagus cuniculus*	OcuZinedin	XP_002722925.1
30	Zinedin	Monkey	Mammal	*Callithrix jacchus*	CjaZinedin	XP_002807900.1
31	Zinedin	Zebra finch	Bird	*Taeniopygia guttata*	TguZinedin	ENSTGUP00000014266
32	Zinedin	Zebrafish	Fish	*Danio rerio*	DreZinedin	XP_691246.3
33	Zinedin like	Tilapia	Fish	*Oreochromis niloticus*	OniZinedin	XP_003450888
34	Unnamed protein	Pufferfish	Fish	*Tetraodon nigroviridis*	TniZinedin	CAF97462

Table 2. Lower eukaryotic organisms studied

SN	Protein	Genera	Scientific name	Abbreviation	Accession no.
1	Cell differentiation and development protein Fsr1/Pro11	Fungi	*Talaromyces stipitatus ATCC 10500*	TstFsr1	XP_002341977.1
2	Striatin Pro11	Fungi	*Verticillium albo-atrum VaMs.102*	ValPRO11	EEY19341.1
3	STRN_SORMA	Fungi	*Sordaria macrospora*	SmaPRO11	Q70M86.1
4	YGD1_SCHPO	Yeast	*Schizosaccharomyces pombe*	SpoWDR	O94560.1
5	STRN_DICDI	Dictiostelium	*Dictyostelium discoideum*	DdiSTRNHomolog	Q54J37.1
6	predicted protein	Fungi	*Nectria haematococca mpVI 77-13-4*	NhaPredicted	XP_003049594.1
7	Pro11	Fungi	*Paracoccidioides brasiliensis Pb01*	PbrPro11	XP_002789630.1
8	WD domain-containing protein	Fungi	*Ajellomyces capsulatus G186AR*	AcaWDD	EEH05528.1
9	WD domain-containing protein G Beta repeat	Fungi	*Coccidioides posadasii C735 delta SOWgp*	CpoWDD	XP_003069696.1
10	hypothetical protein	Fungi	*Trichophyton verrucosum HKI 0517*	TveHypothetical	XP_003018071.1
11	hypothetical protein SS1G_02036	Fungi	*Sclerotinia sclerotiorum 1980*	SscHypothetical	XP_001597840.1
12	hypothetical protein ARB_00865	Fungi	*Arthroderma benhamiae CBS 112371*	AbeHypothetical	XP_003012982.1
13	hypothetical protein BC1G_02788	Fungi	*Botryotinia fuckeliana B05.10*	BfuHypothetical	XP_001558717.1
14	hypothetical protein	Fungi	*Podospora anserina S mat+*	PanHypothetical	XP_001906852.1
15	hypothetical protein An16g01520	Fungi	*Aspergillus niger*	AniPRO11	XP_001397465
16	Conserved Hypothetical	Fungi	*Uncinocarpus reesii 1704*	UreConservedHypo	XP_002544504.1
17	stalk rot protein	Fungi	*Gibberella moniliformis*	Gmostalk rot protein	AAX55652.1
18	Pro11	Fungi	*Neurospora crassa OR74A*	NcrPro11	XP_963602.1
19	Cell differentiation and development protein Fsr1/Pro11	Fungi	*Penicillium marneffei ATCC 18224*	PmaPro11	XP_002151073.1
20	WD domain-containing protein	Fungi	*Ajellomyces dermatitidis SLH14081*	AdeWDD	XP_002626476.1
21	Connector of kinase to AP-1, isoform A	Fly	*Drosophila melanogaster*	DmeCKA	NP_609177.1
22	PREDICTED: similar to striatin,	Hydra	*Hydra magnipapillata*	HmaStrnHomolog1	XP_002158898.1
23	PREDICTED: similar to nuclear autoantigen	Hydra	*Hydra magnipapillata*	HmaStrnHomolog2	XP_002165388.1
24	Calmodulin binding protein 3	Hydra	*Hydra vulgaris*	HvuStrn3	AFA36449.1

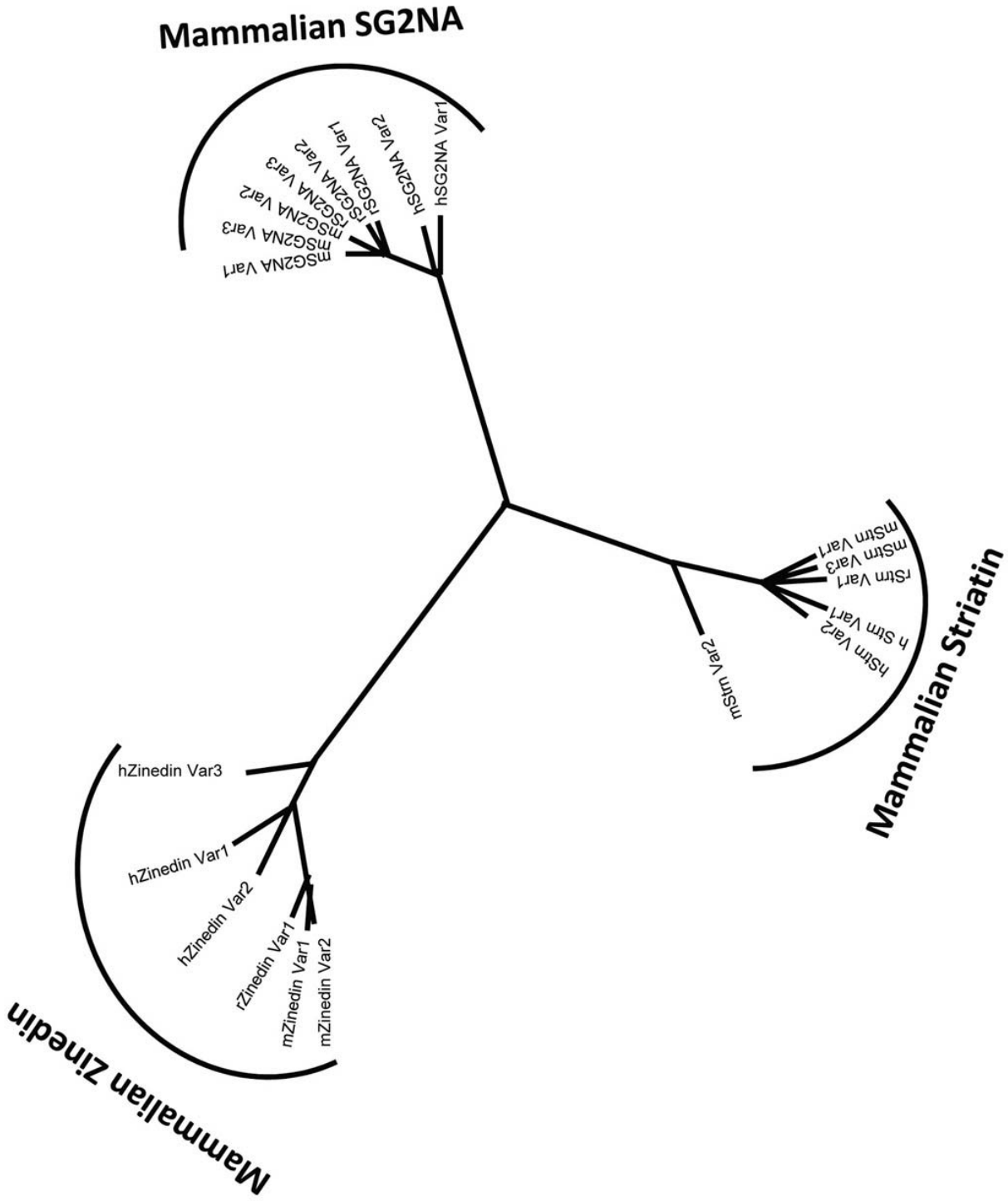


Figure 1. Phylogenetic analysis of striatin, SG2NA, and zinedin from higher eukaryotes. The protein sequences were aligned by using MAFFT program; and gap-removed alignment was used for phylogenetic analysis by Neighbour Joining method. The phylogenetic tree shows three members of the family form three different clusters. The variants of each member are marked by a numerical at the end.

Striatin-like proteins in lower metazoans such as *D. melanogaster* (CKA) and *H. magnipapillata* were closer to their counterparts in higher eukaryotes, whereas the one in *S. pombe* was closer to those present in nonmetazoan eukaryotes. A similar phylogenetic tree was also generated by aligning the homologs of the C-terminal WD-40 repeat

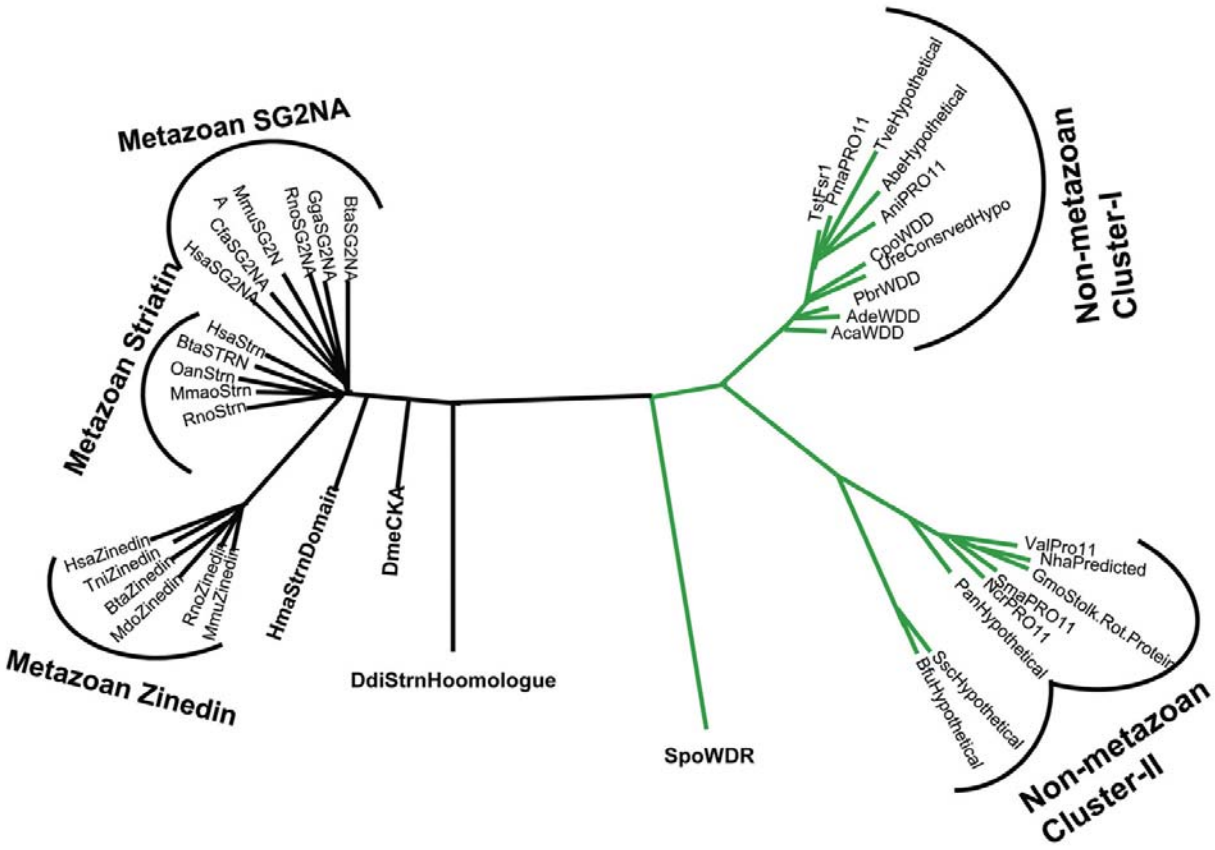


Figure 2. Unrooted phylogenetic tree generated with the N-terminal region (striatin domain) of striatin family of proteins from eukaryotes. The sequences were aligned by MAFFT program; gap-removed alignment was used for generating the phylogenetic tree. Two clusters of yet-unknown identity were seen in lower organisms. In higher organisms, three clusters representing striatin, SG2NA, and zinedin were seen. Striatin homologs from *Schizosaccharomyces pombe* (SpoWDR), *Dycteostelium discoideum* (DdiStrn), *Drosophila melanogaster* (DmeCKA), and *Hydra magnipapillata* (HmaStrnDom) are situated in the linking region between lower to higher eukaryotes.

sequence of striatin. The higher eukaryotic homologs of the C-terminal also formed three distinct subclusters composed of striatin, SG2NA, and zinedin (Figure 3). In this phylogenic tree, Drosophila striatin (CKA) and its homolog in hydra clustered with striatin in higher eukaryotes, whereas their counterpart in *S. pombe* was closer to those in lower eukaryotes.

3.3. Striatin families are characterized by a number of conserved motifs

Striatin family has four distinct domains from the N-terminus to the C-terminus, viz., caveolin-binding domain, coiled-coil domain, calmodulin-binding domain (CaM), and WD-40 repeat domain (12, 15). Alignment of metazoan striatin, SG2NA, and zinedin sequences by using MAFFT and subsequent gap removal showed 11 conserved motifs (eCM I–XI; Figure 4; Table 4). Those motifs were then confirmed using Gblock analysis. Of 11 motifs thus identified, eCM I–III were located in the N-terminus, whereas eCM V–XI were designated at the C-terminus of striatin family members. ScanProsite analysis retrieved specific hits for all conserved motifs, except for eCM IV, which had ≥1000 hit, reflecting its nonspecificity.

3.4. Members of the striatin family have their characteristic motifs

We further examined each members of the striatin family to identify their individual characteristic motifs, if any. Thirteen motifs labeled sCM I–XIII were found to be conserved among the striatin proteins (Figure 5 upper panel, Table 5), of which sCM I and sCM IV were unique to striatin and absent in both SG2NA and zinedin. These two motifs, present in the N-terminus region of the striatin protein, correspond to the regions flanking eCM II of striatin family members. ScanProsite analysis also revealed sCM VII as another striatin-specific motif. In spite of a large part of the motif being conserved among all the striatin family members, unique amino acid residues at the N-terminus region of sCM VII makes it specific to striatin family members. Among the rest of the conserved motifs in striatin, sCM I and sCM IX were also part of eCM I and eCM XI, respectively. The motif sCM VIII, which encompasses the WD-40 repeats, was large and was therefore subdivided into sCM VIIIa–c, of which sCM VIIIb (P-[IL]-Y-T-F-R-[AS]-H-X-G-P-V-L-C-[LV]-[AV]-[MT]) was also present in SG2NA and zinedin; thus, it appears to be conserved among the family members .

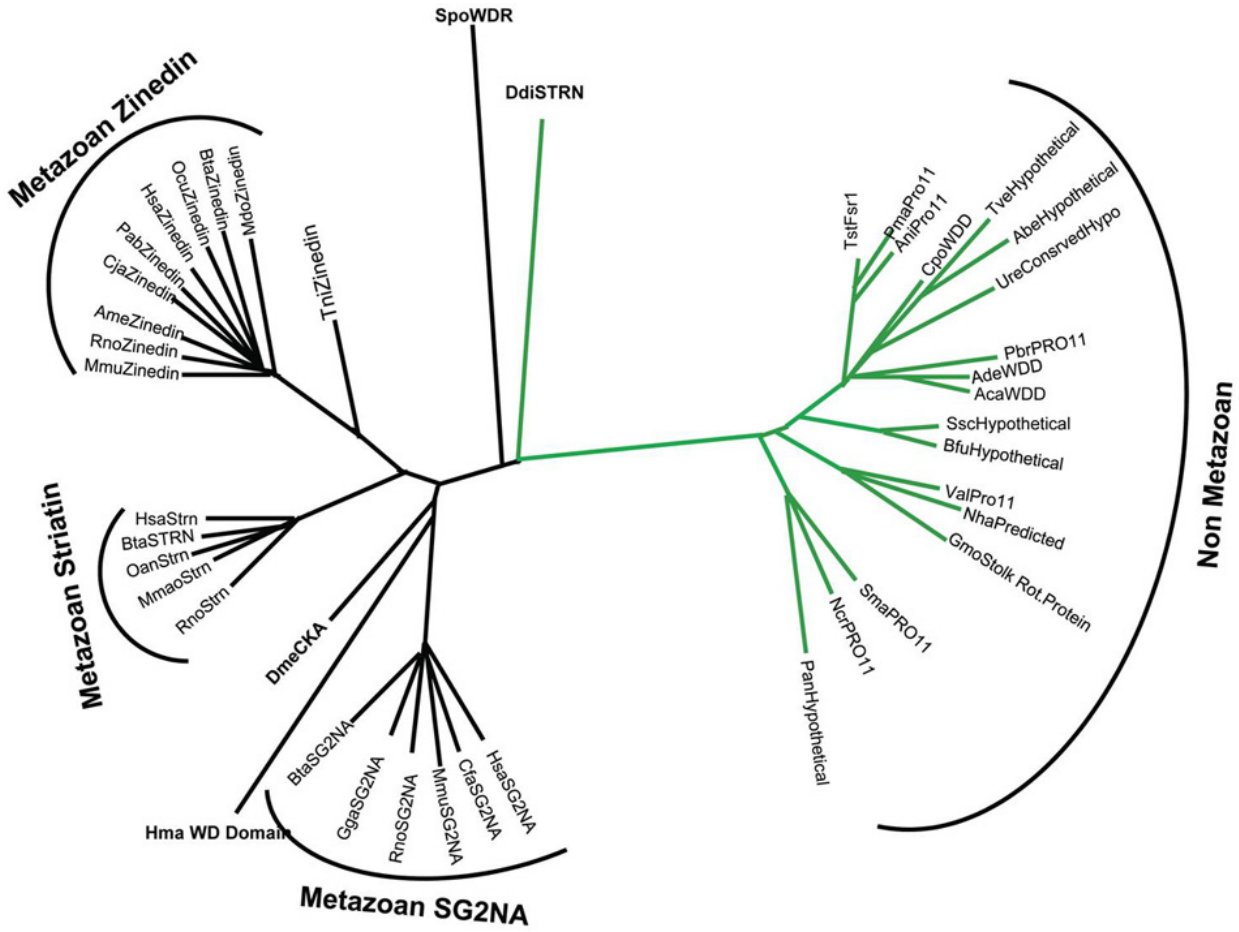


Figure 3. Unrooted phylogenetic tree generated with C-terminal region (WD-40 domain) of striatin family of proteins from eukaryotes. The sequences were aligned by MAFFT program; gap-removed alignment was used for generating the phylogenetic tree. Lower eukaryotes formed a single cluster, whereas higher eukaryotes formed three clusters. Homologs from *Schizosaccharomyces pombe* (SpoWDR) and *Dycteostelium discoideum* (DdiStrn) are situated at the linking position of lower and higher eukaryotes. Striatin homolog from *Drosophila melanogaster* CKA (DmeCKA) and *Hydra magnipapillata* (HmaStrnDom) are situated in the SG2NA cluster.

Using similar methodology based on homology to delineate conserved motif patterns in metazoan SG2NA homologs revealed at least 10 conserved motifs (Figure 5 middle panel, Table 6). sgCM II was identified as unique motif of SG2NA and corresponded to the region just upstream to that of eCM I. The unique amino acid residues at its C-terminus end made it specific to SG2NA. As in the case of striatin, the conserved motif sgCM VIII encompassing WD-40 repeats was subdivided into four smaller motifs, viz., sgCM VIIIa–d. Of these, sgCM VIIId was found to be specific to SG2NA. ScanProsite analysis of this motif revealed its matches in nonmetazoan (Nostoc species) and bacterial species (data not shown). The conserved motif sgCM IX encompasses eCM XI, while the conserved motif sgCM X is downstream to eCM XI. Although sgCM X is well conserved in other striatin family members, it is not reflected as metazoan-conserved motifs, as this region is absent in certain organisms such as Callithrix jacchus (in zinedin) and T. nigroviridis (in unnamed protein), resulting in gaps. This was thus deleted in the final analysis.

Similarly, sequence conservation analysis of metazoan zinedin homologs revealed at least nine conserved motifs (Figure 5 lower panel, Table 7), of which zCM VII and zCM IX were large and were thus subdivided into smaller motifs. In addition, zCM VII was shown to encompass the WD-40 repeat region. zCM II and zCM VI, consisting of 25 and 37 amino acid residues, respectively, retrieved hits to specific zinedin family members in ScanProsite analysis, thus rendering them unique to zinedin. zCM II corresponded to the region upstream to eCM II whereas zCM VI corresponded to that of eCM IV. Although the core region of zCM VI is homologous to other striatin family members, its specificity can be attributed to the flanking amino acid sequence unique to zinedin only.

In the past decades, researchers showed that multidomain structure of striatin family members is well defined. Determining the amino acid sequence of each of these domains has also been well established (15).

Table 3. Prokaryotic organisms studied

SN	Protein name	Genera	Scientific name	Abbreviation	Accession no.
1	Metallophosphoesterase	Bacteria	*Thermomonospora curvata DSM 43183*	TcuMetalloprotease	YP_003300732.1
2	Hypothetical protein	Bacteria	*Microcoleus chthonoplastes PCC 7420*	MchHypthetical	ZP_05024620.1
3	Peptidase C14, caspase catalytic subunit p20	Bacteria	*Cyanothece sp. CCY0110*	CspPeptidase	ZP_01727437.1
4	WD repeat-containing protein	Bacteria	*Nostoc sp. PCC 7120*	NspWD repeat	NP_486840.1
5	WD repeat protein	Bacteria	*Streptomyces sviceus ATCC 29083*	SsvWD repeat	ZP_06922085.1
6	WD-40 repeat protein	Bacteria	*Oscillatoria sp. PCC 6506*	Osp WD-40 repeat	ZP_07109766.1
7	Putative regulatory protein	Bacteria	*Mycobacterium marinum M*	MmaPutative Regulatory	YP_001851685.1
8	Fis family transcriptional regulator	Bacteria	*Anabaena variabilis ATCC 29413*	AvaFis family	YP_323138.1
9	WD repeat-containing protein	Bacteria	*Gloeobacter violaceus PCC 7421*	GviWD repeat	NP_924121.1
10	Myosin heavy-chain kinase	Bacteria	*Haliscomenobacter hydrossis DSM 1100*	HhyMyosin Heavy chain kinase	YP_004451363.1
11	Metallophosphoesterase	Bacteria	*Thermomonospora curvata DSM 43183*	TcuMetalloprotease	YP_003300732.1
12	WD repeat-containing protein	Archaea	*Methanosarcina barkeri str. Fusaro*	MbaWD repeat	YP_304107.1
13	WD-40 repeat-containing protein	Archaea	*Methanosarcina acetivorans C2A*	MacWD-40 repeat	NP_617428.1
14	Hypothetical protein MCP_1127	Archaea	*Methanocella paludicola SANAE*	MpaHypothetical	YP_003356182.1
15	WD-40 repeat-containing protein	Archaea	*Methanosarcina barkeri str. Fusaro*	MbaWD-40	YP_306342.1
16	WD-domain-containing protein	Archaea	*Methanosarcina acetivorans C2A*	MacWD domain	NP_617481.1
17	Hypothetical protein Mbar_A0545	Archaea	*Methanosarcina barkeri str. Fusaro*	MbaHypothetical	YP_304106.1

Therefore, we conducted positional distribution analysis to determine whether the conserved motifs identified so far in our analysis represent any of these signature domains among striatin family members. Our results showed that both coiled-coil and CAM domains were well conserved among SG2NA counterparts throughout metazoans, represented by sgCM I and sgCM III. Although caveolin-binding and coiled-coil domains are well represented in almost all striatin family members, these domains were missing in few striatin and zinedin homologs among metazoans (e.g., striatin in *Pongo abelii* and zinedin in *Oryctolagus cuniculus*). This finding suggests that CaM domain—represented by eCM I (sgCM III in SG2NA)—is the most conserved domain of the three N-terminus domains among metazoan striatin family members. Among metazoan striatin counterparts, calmodulin-binding domain is represented by sCM I, whereas in zinedin, it is represented by zCM I. In both calmodulin-binding and WD-40 repeat domains, striatin and zinedin have four conserved motifs each (i.e., sCM II–V and zCM II–V, respectively), whereas SG2NA isoforms have between two to four conserved motifs, suggesting variation because of alternative splicing of exons 8 and 9. Although it is not highlighted in the Gblock analysis (due to gaps because of splicing in SG2NA), the amino acid sequence encoded by exon 8 in SG2NA is highly identical to that in striatin and is completely absent in zinedin. This is clearly observed in homology match between full-length SG2NA and striatin protein sequences. Although no splicing has been reported in zinedin at exon 8, it is shown to encode for the zinedin-specific amino acid sequence, that is, zCM VI, with least homology to that encoded by the same exon in SG2NA and striatin.

3.5. Striatin homologs in lower eukaryotes also contain conserved motifs

In recent years, *in silico* and motif-deletion studies in *Fusarium verticillioides FSR1* (27-28) and *A. nidulans StrA* (29), the non-metazoans striatin like proteins, reveal that they retain all the four domains similar to that in metazoan striatin counterparts. Functional complementation analyses, using various truncated versions, also demonstrated the role of *StrA* in sexual development in *A. nidulans* and of coiled-coil domain of *FSR1* in various protein–protein interactions in mediating *F. verticillioides* virulence. By using available amino acid sequence alignment tools, we sought to identify the motif pattern among lower eukaryotic striatin homologs and their conservation in higher eukaryotic counterparts. In this aspect, striatin homologs from nonmetazoans were retrieved from NCBI database by using striatin of lower eukaryotes such as *Drosophila* and hydra as query. MAFFT alignment and Gblock analyses of the sequences thus obtained led to the identification of 13 conserved motifs (Figure 6; Table 8). As leCM I and XII were longer than 200 residues, they were subdivided into two submotifs each; neither showed any significant sequence homology with the motifs conserved in their metazoan counterparts. ScanProsite showed that leCM I and IV are present in both metazoan and nonmetazoan homologs of striatin, where leCM I encompassed all the three N-terminus domains, that is, caveolin-binding, coiled-coil, and CAM domain. leCM II and VIII were considered nonspecific as they were present in many unrelated proteins in the database. leCM VII was found in SG2NA as well as other WD-40 repeat–containing proteins. The remaining motifs were unique to striatin homologs in nonmetazoan organisms.

3.6. Striatin family members have evolved from prokaryotes

Comparison of the human striatin sequence with the bacterial and archaeal proteomes showed that the homology was restricted to the C-terminal domains of the bacterial and archaeal proteins. Eleven best-matched sequences from bacteria and 6 best-matched sequences from archaeal proteomes were taken for further analysis (Table 3). These sequences were aligned using MAFFT, and gaps were removed. It was observed that these sequences were conserved only in their C-terminal domain. Within the C-terminal domain, four conserved motifs were identified (Table 9) and named pCM I–IV. Analysis of pCM I using ScanProsite led to the identification of metazoan SG2NA, but not striatin and zinedin. pCM II matched with more than 10,000 hits and was thus considered nonspecific. Analysis of pCM III and IV identified WD-repeat–containing proteins, but none

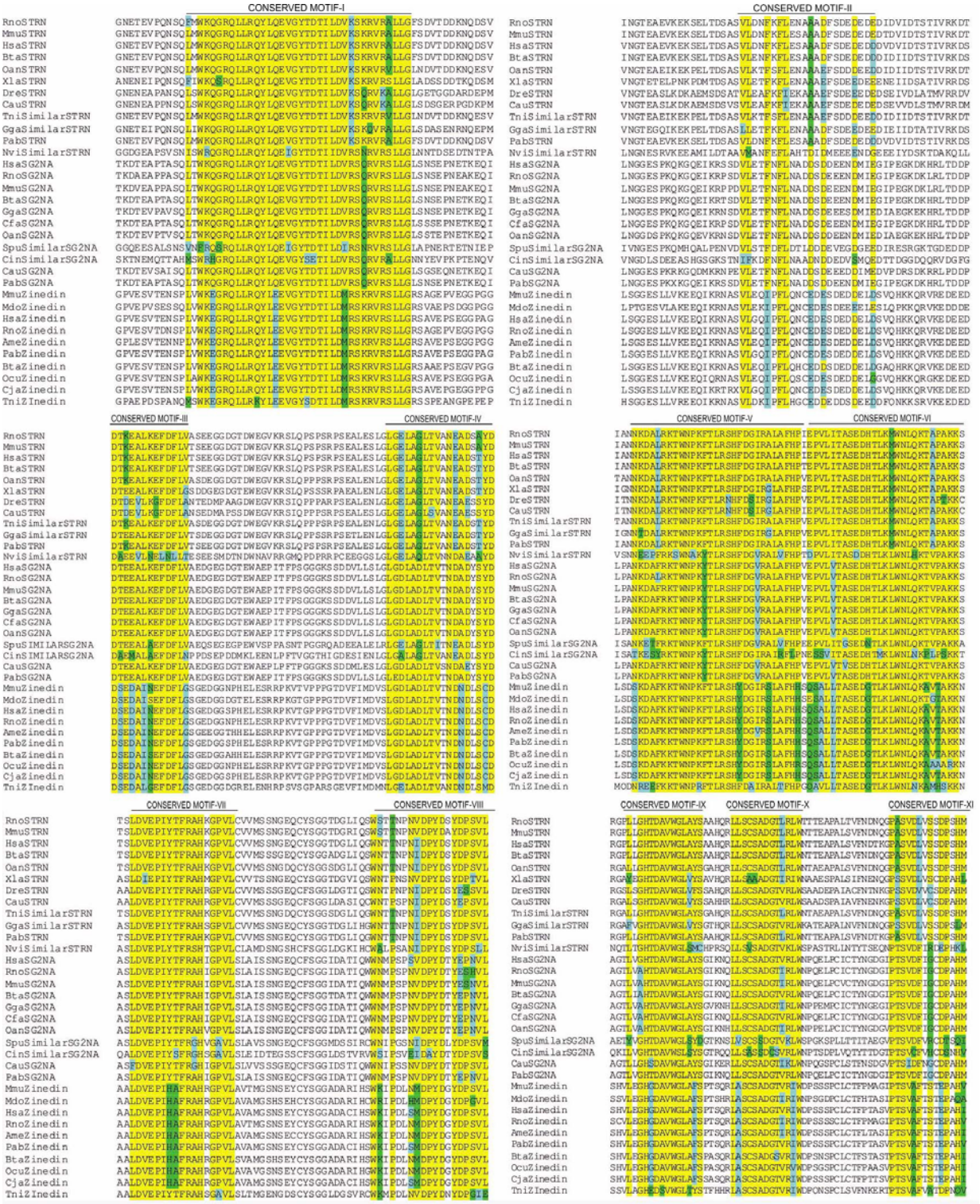


Figure 4. Motif conservation in striatin family members in higher eukaryotes. Sequences of striatin family of proteins from higher eukaryotes were aligned by MAFFT program, and gaps were removed. Eleven conserved motifs were identified that were further verified by Gblock software.

belonged to the striatin family. Data suggest that only pCM I is conserved between the SG2NA and the bacterial WD-40 proteins. To understand which of the three members of the family evolved first, we scored for the presence and

Table 4. Conserved motifs (CMs) in striatin family of proteins in higher eukaryotes

Conserved motifs of higher eukaryotes	Motif sequences/pattern	Minimum element
eCM I	[FLVM]-X-[WF]-[KR]-[QHE]-[GS]-R-Q-L-L-R-[QK]-Y-L-[QE]-E-[IV]-G-Y-[ST]-[DE]-T-I-L-D-[VIM]-[KR]-S-[KQN]-[RQ]-V-[KR]-[AVS]-L-L-G	[FLVM]-X-[WF]-[KR]-[QHE]-[GS]-R-Q-L-L-R
eCM II	[VI]-[LMF]-[DEKASG]-X-[FI]-X-F-[LI]-[EANDQH]-X-[ATDLC]-[ADE]-[AISDN]-[DE]-[FMDES]-[SEVD]-[DE]-[EDN]-[DES]-[EDNMGI]-[DEIQML]-[DEG]	[VI]-[LMF]-[DEKASG]-X-[FI]-X-F-[LI]-[EANDQH]-X-[ATDLC]-[ADE]-[AISDN]-[DE]-[FMDES]
eCM III	D-[TAS]-[KED]-[EMD]-[AV]-[LI]-[KNAS]-[EG]-[FL]-[DN]-[FL]-L-[VGATN]	D-[TAS]-[KED]-[EMD]-[AV]-[LI]-[KNAS]-[EG]-[FL]-[DN]
eCM IV	L-G-[DEA]-L-A-[GD]-L-[TS]-V-[ANTS]-N-[DE]-[AN]-[DE]-[SAYL]-[ATS]-[YC]-D	L-G-[DEA]-L-A-[GD]-L-[TS]-V
eCM V	[NTS]-[KTE]-[DE]-[APTSE]-[LFY]-R-K-[ST]-W-N-[PA]-[KR]-[FY]-T-L-R-[SN]-H-[FY]-D-[GA]-[IV]-R-[AGS]-L-[AVRT]-F-[HL]-[PH]	[NTS]-[KTE]-[DE]-[APTSE]-[LFY]-R-K-[ST]-W
eCM VI	[EDQ]-[PSA]-[VSA]-[LV]-[IV]-T-[AGV]-S-[DE]-D-[HNG]-T-[LM]-K-[LM]-W-N-L-[QHN]-K-[TPA]-[AVLGM]-[PTA]-[STA]-[KR]-K	[EDQ]-[PSA]-[VSA]-[LV]-[IV]-T-[AGV]-S-[DE]-D
eCM VII	[LF]-D-[IV]-E-P-I-[YH]-[STA]-F-R-[AG]-H-X-G-[PA]-V-L	[LF]-D-[IV]-E-P-I-[YH]-[STA]
eCM VIII	W-[SNAKR]-[TLMI]-[PT]-[NSD]-[PSAVL]-[NSEH]-[AVM]-D-[PA]-Y-D-[STAGN]-Y-[DE]-[PS]-[STNHG]-[IVL]-[LMSE]	W-[SNAKR]-[TLMI]-[PT]-[NSD]-[PSAVL]-[NSEH]-[AVM]-D-[PA]
eCM IX	[LYF]-[LESVTA]-[GA]-H-[STGE]-D-[AS]-V-W-G-L-[AVST]-[YMF]-[SD]	[LYF]-[LESVTA]-[GA]-H-[STGE]-D-[AS]-V-W-G
eCM X	L-[LVA]-S-[CAV]-[SA]-[SA]-D-[GC]-[ST]-[VIL]-[KR]-[LI]-W	Nonspecific
eCM XI	P-[STA]-S-[IV]-[DCA]-[FL]-[VINT]-[SCRGH]-[SCDT]-[DE]-[STP]-[SAHN]-[HQL]-[MLIVA]	P-[STA]-S-[IV]-[DCA]-[FL]-[VINT]-[SCRGH]-[SCDT]-[DE]

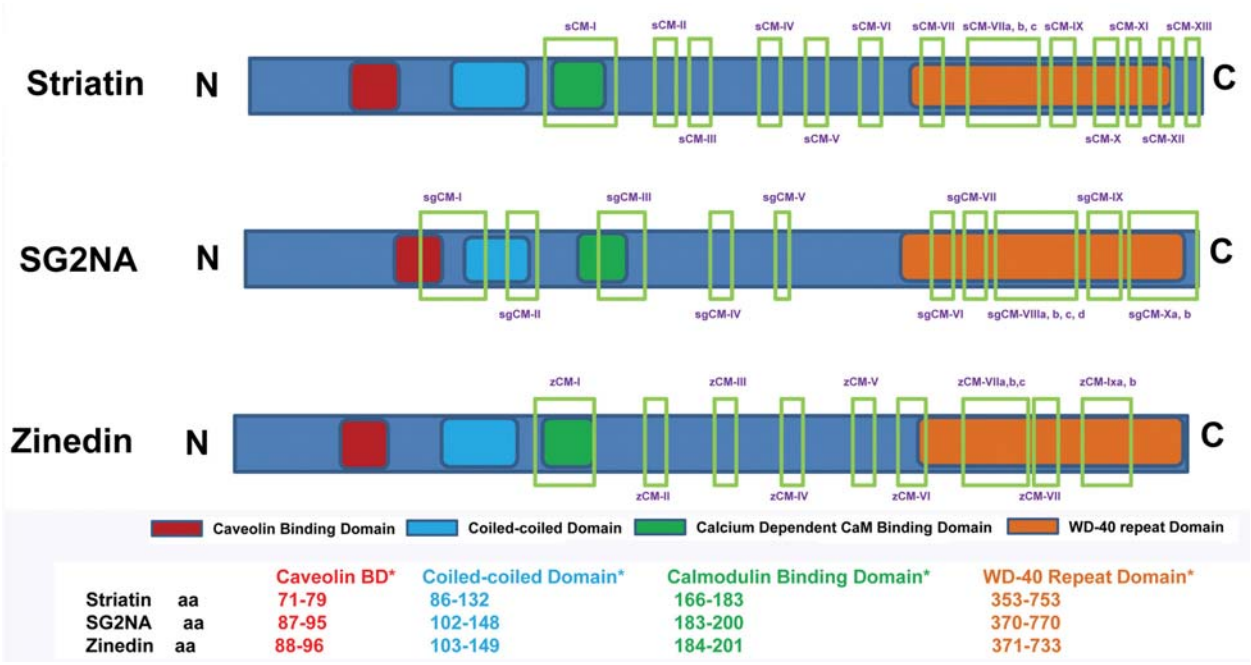
Table 5. Conserved motifs (CMs) in higher eukaryotes, specifically in striatin

Conserved motifs present only in striatin	Motif sequences/pattern	Minimum residues to identify striatin
sCM I	S-[QN]-[FLI]-[MSI]-W-[KR]-Q-[GS]-R-Q-L-L-R-Q-Y-L-Q-E-[VI]-G-Y-T-D-T-I-[LI]-D-V-K-S-[KQ]-[RQ]-V-[KR]-[VAS]-L-L-G-[FL]-[GANS]	S-[QN]-[FLI]-[MSI]-W-[KR]-Q-[GS]
sCM II	[TPHN]-[AVLTS]-[MVS]-I-[GA]-[KE]-X-[EDM]-[MLI]-[STL]-D-[ST]	[TPHN]-[AVLTS]-[MVS]-I-[GA]-[KE]-X-[EDM]-[MLI]-[STL]-D-[ST]
sCM III	[AV]-[STA]-[LV]-[LM]-[DEA]-[TNA]-F-[SKE]-F-[IL]-[EK]-[KNSH]-[AT]-[AD]-[AT]-[DE]-[FM]-[SE]-[DE]-[DE]-[DE]-[DEN]-[DE]-[DEG]-[DE]-X-[DE]	[AV]-[STA]-[LV]-[LM]-[DEA]-[TNA]-F-[SKE]-F-[IL]
sCM IV	[IV]-[IVY]-D-[STL]-[SAK]-T-[IMD]-[VA]-[RK]-[KRQ]	[IV]-[IVY]-D-[STL]-[SAK]-T-[IMD]-[VA]-[RK]
sCM V	[TA]-[KDE]-E-[VA]-L-[KN]-[EG]-[FL]-[DN]-[FL]-L-[VATG]-X-[SDT]-[DE]-[DGE]	[TA]-[KDE]-E-[VA]-L-[KN]-[EG]-[FL]-[DN]-[FL]-L-[VATG]-X-[SDT]
sCM VI	[GAMP]-[DAP]-[TASG]-[NGST]-[DE]-W-[NE]-[AK]-[VE]-[RED]-[RQ]-[GSC]-[MPL]-X-[PCST]-[DE]	[GAMP]-[DAP]-[TASG]-[NGST]-[DE]-W-[NE]-[AK]-[VE]-[RED]
sCM VII	E-[EAT]-[GL]-[GE]-[SN]-[SE]-L-G-L-G-E-L-A-[QG]-L-[ST]-V-[AN]-N-[DE]-A-[DE]-[SA]	E-[ATE]-[LG]-[EG]-[NS]-[ES]-L-G-L
sCM VIIIa	[STAG]-N-N-[KET]-[DE]-[PA]-[FL]-R-K-[ST]-W-N-[PA]-K-[FY]-T-L-R-[SN]-H-F-D-[SG]-[IV]-R-[AG]-L-[AV]-F-H-P-X-[DE]-P-V-L-I-T-A-S-[DE]-D-H-T-L-K-[ML]W-N-L-[HQ]-K-T-[AV]-P-[AT]-K-K-[SC]-[AT]-[SA]-L-D-[IV]-E	[STAG]-N-N-[KET]-[DE]-[PA]-[FL]-R-K-[ST]
sCM VIIIb	P-[IL]-Y-T-F-R-[AS]-H-X-G-P-V-L-C-[LV]-[AV]-[MT]-[SD]-S-[NS]-G-[SED]-[HQ]-C-[FY]-S-G-G-X-D-[GA]-X-I-[QH]-X-W-[NAS]-[LT]-[PT]-[SN]-[IV]-D-P-Y-D-[SA]-Y-[DE]-[PS]-[ST]-[LV]-L-[RN]-[GQ]-X-[LYF]-X-G-H-T-D	P-[IL]-Y-T-F-R-[AS]-H-X-G-P-V-L-C-[LV]-[AV]-[MT]-[SD]
sCM VIIIc	A-V-W-G-L-[SVA]-[YFM]-[SC]-X-[PAVT]-[HR]-[SQH]-[RQ]-L-L-S-[VCA]-[SA]-A-D-G-T-[IVL]-[RK]-[IL]-W-[SNKT]-[PAT]-X-[SDE]-X-[TESA]-[PR]-[LA]-[IL]-[NAST]-[TCVI]-[YF]-[TN]-[STDE]-[NE]-[QK]-[DE]-[LRM]-[GN]	A-V-W-G-L-[SVA]-[YFM]-[SC]-X-[PAVT]-[HR]
sCM IX	P-[TSA]-S-V-D-[FVL]-[IV]-[RCS]-[DSC]-[DE]-P-[HAS]-[HKL]-[LM]	P-[TSA]-S-V-D-[FVL]-[IV]-[RCS]-[DSC]
sCM X	[IL]-[FY]-[ND]-[MA]-E-T-X-[QS]-[ILKR]-X-[LS]	[IL]-[FY]-[ND]-[MA]-E-T-X-[QS]-[ILKR]-X-[LS]
sCM XI	[IT]-L-P-[IL]-[VTS]-[VI]-[AT]-A-[NQH]-E-D-R-H-I-[KR]-F-[YF]	[IT]-L-P-[IL]-[VTS]-[VI]-[AT]-A-[NQH]-E-D-R-H-I-[KR]-F-[YF]
sCM XII	[NR]-[MAL]-[DKE]-[NPS]-[KA]-[ST]-[RC]-[IVS]-[QS]-[ER]-[IPF]-T-[GA]-[HR]-[RS]-[SK]-[RK]-[FRS]	[NR]-[MAL]-[DKE]-[NPS]-[KA]-[ST]-[RC]-[IVS]-[QS]-[ER]-[IPF]-T-[GA]-[HR]-[RS]-[SK]-[RK]-[FRS]
sCM XIII	P-[GST]-[RK]-[PTC]-[FRY]-[IS]-[APG]-[SK]-[AC]-[GS]	P-[GST]-[RK]-[PTC]-[FRY]-[IS]-[APG]-[SK]-[AC]-[GS]

absence of the conserved motifs by using ScanProsite. It was found that the sgCM VIIId was present in two prokaryotes—Nostoc and Mycobacterium marinum (data not shown). None of the conserved motifs present in striatin or zinedin identified any prokaryotic organisms by ScanProsite analysis (data not shown). Similarly, pCM I was present not only in S. pombe but also in metazoan SG2NA (data not shown). The identification of conserved motifs of lower eukaryotes (leCM VII) and that of prokaryotes (pCM I), in SG2NA and vice versa shows that among the three striatin family members SG2NA is the most conserved protein in both eukaryotes and prokaryotes.

4. DISCUSSION

In the post genomic era, rapid development of bioinformatic tools has accelerated our ability to predict structure–function relationship of proteins from annotated sequences. Comparative functional analysis through sequence alignment is easier in case of proteins with similar sequences. However, studying those with lesser sequence similarities (~25% identity) but analogous functions leads in to "twilight zone," requiring more sophisticated tools (30). Since WD-40 repeat proteins are highly pervasive in occurrence and versatile in functions, studying their evolution is an opportunity as well as a



	Caveolin BD*	Coiled-coiled Domain*	Calmodulin Binding Domain*	WD-40 Repeat Domain*
Striatin aa	71-79	86-132	166-183	353-753
SG2NA aa	87-95	102-148	183-200	370-770
Zinedin aa	88-96	103-149	184-201	371-733

Figure 5. Distribution of conserved motifs in different domains of striatin family members. Distribution of conserved motifs in striatin, SG2NA, and zinedin were analyzed manually and are shown schematically. The four domains of striatin homologs previously marked by Castets *et al*. (15) are indicated by asterisks (*).

challenge. Many WD-40 proteins have lower frequency of sequence conservation and are thus difficult to detect in the genome (3, 5). On the other hand, since in the course of evolution they have progressively acquired various other domains conferring functions as diverse as ubiquitylation and protein degradation, vesicle coating, phospholipids binding, and so forth, they also provide an opportunity to probe the evolutionary road map in general and that of metazoans in particular. We thus embarked upon the evolutionary analysis of the striatin subfamily of WD-40 repeat proteins.

Although the three members of the family have been identified for quite some time now, little is known about their roles in metazoan biology. They are primarily membrane associated and presumably function as membrane scaffold (12). In addition, due to the existence of multiple splice variants, they are anticipated to have diverse interacting partners (21). Thus, we argued that analyzing their evolutionary path might provide an insight into their functions, especially in the context of the evolution of metazoan interactomes.

The patterns identified, based on homology, depict diversification of striatin family members into newer variants during evolution with certain motifs being well conserved. Striatin and SG2NA are widely present in almost all metazoans. Mammals have all three members, including zinedin, but platypus has only striatin and SG2NA. Similarly, organisms ranging from hydra to Drosophila have only one orthologue each. Collectively, it appears that striatin family members have adapted to the increasing evolutionary complexity by evolving newer family members. Also, in the phylogenic clusters, each of the three family members maintained distinctive identity and thus appears to be functionally discrete. As expected, all the variants of striatin, zinedin, and SG2NA clustered together, suggesting their functional relatedness.

Of the 11 conserved motifs identified in all three members of the family, R-Q-L-L-R (eCM I) is a variant of the canonical IQ motif (IQXXXRGXXXR), found in proteins with Ca^{++}–independent calmodulin-binding activities. However, IQ motif has also been found in a novel nuclear protein IQD1 from Arabidopsis, which binds Calmodulin in a Ca^{++}–dependent manner (31). Noticeably, calmodulin-binding activity of both SG2NA and striatin are Ca^{++} dependent (15), while upon serum stimulation, SG2NA transiently localizes to the nucleus (unpublished results; in contrary to the observation by Muro *et al.* [1995], who showed that SG2NA is a resident nuclear protein). Since striatin family members are absent in plants, it would be interesting to see whether there are any other commonalities between SG2NA and IQD1. The significance of the other invariant motif V-W-G-L (eCM IX), found in all three members of the family, is not clear as yet. However, this motif and six others (eCM V-XI) in the WD-40 repeat domain suggest their roles in specifying the β-propeller structures. During evolution, WD-40 domains have diversified by gene duplication and recombination, thereby providing interactive platforms for multiple proteins. It is likely that these motifs define the functional relatedness as well as specificities of the WD-40 domains of the striatin subfamily (3-4). Noticeably, caveolin-binding and coiled-coil domains were found in striatin and zinedin homologs of almost all metazoans

Table 6. Conserved motifs (CMs) in higher eukaryotes, specifically in SG2NA

Conserved motifs present only in SG2NA	Motif sequences/pattern	Minimum residues to identify SG2NA
sgCM I	[VL]-E-[KR]-[AE]-[EQ]-L-[QK]-A-R-I-A-F-L-Q-G-E-R-K-G-Q-E-N-L-K-[RHNK]-D-L-[IV]-R-R-I-K-M-L-E-[FY]-A-L-K-Q-E-R-[AT]-K-[VY]-[HQ]-K-L-K-[FSY]-G-[ST]-[DE]	[VL]-E-[KR]-[AE]-[EQ]-L-[QK]-A-R-I-A-F-L-Q-G-E-R
sgCM II	[MKQ]-[EDG]-[QED]-[KLM]-[HK]-[NAM]-[EP]-[TS]-[NF]-[ED]-[EGS]	[MKQ]-[EDG]-[QED]-[KLM]-[HK]-[NAM]-[EP]-[TS]-[NF]-[ED]
sgCM III	[MVL]-[SNT]-[WF]-[RK]-[QH]-[GS]-R-Q-L-L-R-Q-Y-L-Q-E-[VI]-G-Y-[ST]-[ED]-T-I-L-D-[IV]-R-S-[QN]-R-V-R-[AS]-L-L-G-[NL]-[NAS]-[YPGNS]-[ENST]-[VE]-[PRQ]-[KTN]-[PG]-[TPS]	[MVL]-[SNT]-[WF]-[RK]-[QH]-[GS]-R-Q-L-L-R-Q-Y-L-Q-E-[VI]-G-Y-[ST]-[ED]-T-I-L-D-[IV]-R
sgCM IV	[IV]-[FL]-[KSE]-[DT]-F-[ND]-F-L-[NSE]-[ADN]-[AT]-[DL]-[ND]-[DS]-[ED]-[DGE]	[IV]-[FL]-[KSE]-[DT]-F-[ND]-F-L-[NSE]-[ADN]
sgCM V	D-[AT]-E-[ME]-A-L-[AK]-E-F-D-F-L-[VN]	D-[AT]-E-[ME]-A-L-[AK]-E-F
sgCM VI	[ESA]-[EP]-[LAI]-[PST]-[FN]-[TP]-[VPS]-G-G-[GST]	[ESA]-[EP]-[LAI]-[PST]-[FN]-[TP]-[VPS]-G-G-[GST]
sgCM VII	D-[ED]-[ESMV]-[IAL]-[EL]-[NES]-[ATV]-L-[GR]-L-[GE]-[AED]-L-A-[GD]-L-T-[IV]-[AST]-N-[ED]-[AE]-[DA]	D-[ED]-[ESMV]-[IAL]-[EL]-[NES]-[ATV]-L-[GR]-L-[GE]
sgCM VIIIa	[LY]-S-Y-D-[IL]-[SP]-A-[TN]-K-[ED]-[STA]-[YFL]-R-K-T-W-N-P-K-[FY]-T-L-R-S-H-F-D-G-[IV]-R-A-[IL]-[RVA]-F-[LH]-P-[NEV]-E-[SP]-[SV]-[VL]-[IV]-T-[AGS]-S-E-D-[HN]-T-[ML]-K-L-W-N-L-[NQ]-K-[PT]-[LV]-P-[SA]-K	[LY]-S-Y-D-[IL]-[SP]-A-[TN]-K
sgCM VIIIb	K-[TAS]-[AQ]-[AS]-[LF]-D-V-E-P-I-Y-[ST]-F-R-[GA]-H-[SVI]-G-[PA]-V-L-S-L-[EAV]-[IV]-[SD]-[TS]-[ENS]-G-[SE]-[SQ]-C-F-S-G-G-[LMIT]-D-[SA]-[ST]-[VI]-[RQ]-[VCW]-W-[SN]-[IM]-P-[SG]-[VSP]-[SEN]-[IV]-D-[AP]-Y	K-[TAS]-[AQ]-[AS]-[LF]-D-V-E-P-I
sgCM VIIIc	D-[LT]-Y-[DE]-[PS]-[SNH]-V-[SML]-[QA]-[KEG]-[IT]-[LY]-[LVI]-[GA]-H-[ST]-D-A-V-W-[GS]-L-[SA]-Y-[SD]-G-[TIV]-[KR]-[QNE]-[QSRH]-L-[LV]-S-[AC]-S-[SA]-D-[CG]-[ST]-[IV]-[RK]-L-W-[SN]-P-[TGQP]-[SKGE]-[AK]	D-[LT]-Y-[DE]-[PS]-[SNH]-V-[SML]-[QA]
sgCM VIIId	[DSML]-P-[CL]-[VLI]-[STQC]-T-[YIF]-[TN]-[TASG]-[DE]	[DSML]-P-[CL]-[VLI]-[STQC]-T-[YIF]-[TN]-[TASG]
sgCM IX	G-[TVI]-P-T-S-[IV]-[CD]-F-[INV]-[RGH]-C-D-[STP]-[NSA]-[HQ]-[IVM]-[IV]-[VAT]-[SG]-[HFY]-[TSN]-[DST]-[GA]-[KDS]-X-[NAV]-[IV]-Y-D-L-E-T-[GAS]-[QK]-[IKPS]-[SVL]-[IMV]-[VITM]-[LF]-[DST]-[SAE]-[RQGH]	G-[TVI]-P-T-S-[IV]-[CD]-F-[INV]-[RGH]-C
sgCM Xa	[HQS]-I-N-[CSKR]-V-[VAS]-[ST]-H-P-T-[ML]-P-[IV]-T-[MIV]-T-A-H-[DE]-D-[RK]-H-I-[KR]-F-[YF]-D-N-[TNK]-[ST]-G-K-[MLAT]-[IV]-H-[AS]-M-V-A-H-L-D-[SA]-V-T-[SC]-L-[SA]-[IV]-D-P-N-G-[LI]-Y-L-[LM]-S-G-S-H-D	[HQS]-I-N-[CSKR]-V-[VAS]-[ST]-H-P
sgCM Xb	[SC]-S-I-R-L-W-N-[LV]-D-S-K-T-C-[IV]-Q-E-[LIV]-T-[SA]-H-R-K-K-[FSL]-D-E-[SA]-I-[YF]-[DS]-V-A-F-H-[PS]-[SF]-[KA]-[TPA]-Y-I-[AG]-S-[AG]-G-A-D-A-L-A-K-V-F-V	[SC]-S-I-R-L-W-N-[LV]-D

Table 7. Conserved motifs (CMs) in higher eukaryotes, specifically in zinedin

Conserved motifs present only in zinedin	Pattern	Minimum residues to identify zinedin
zCM I	[EM]-[QPA]-V-[PSA]-N-G-P-[VAL]-E-[SP]-[DV]-[ST]-[LE]-[PED]-[ASN]-[SN]-[QP]-[LM]-[SVA]-W-K-E-G-R-[KQ]-L-L-R-[KQ]	[EM]-[QPA]-V-[PSA]-N-G-P-[VAL]-E-[DV]
zCM II	[GVE]-P-[GAPE]-[GAP]-[LH]-[SP]-[GST]-[GS]-E-S-[LV]-L-[VA]-[KR]-Q-I-E-E-Q-I-[KQ]-R-[NT]-[AR]-[ASG]	[GVE]-P-[GAPE]-[GAP]-[LH]-[SP]-[GST]-[GS]-E-S
zCM III	[KER]-[DES]-[GS]-[RTK]-E-R-L-[GSA]-[GSA]-X-V-L-[EGD]-[QK]-I-P-F-L-[QH]-[NHSG]-C-E-D	[KER]-[DES]-[GS]-[RTK]-E-R-L-[GSA]-[GSA]-X-V-L
zCM IV	[QK]-[HPN]-[KR]-K-[QS]-R-[VG]-[KR]-[LV]-[PST]-[SR]-[KG]-[APN]-L-[VA]-P-E	[QK]-[HPN]-[KR]-K-[QS]-R-[VG]-[KR]-[LV]
zCM V	[ED]-E-[DE]-[DE]-[DE]-[DE]-D-S-E-D-A-[IL]-[NS]-E-F-D-F-L-G-S-G-E-[DE]-G-E-[VG]-X-[PAHG]-[DE]-[PT]-R-[RI]-[CS]	[ED]-E-[DE]-[DE]-[DE]-[DE]-D-S-E-D-A-[IL]
zCM VI	G-X-[PGH]-[HT]-E-L-E-[SN]-[RH]-R-[VN]-K-L-Q-G-[IM]-[LM]-[AS]-D-[LF]-[RP]-[DS]-[VK]-[DP]-[GA]-[LK]-P-P-[KS]-[VG]-[TS]-[VG]-P-[PA]-[PR]-[GS]-[GT]	G-X-[PGH]-[HT]-E-L-E-[SN]-[RH]-R
zCM VIIa	D-V-F-I-M-D-[TA]-[IV]-G-G-G-[DE]-[VM]-[SN]-L-G-[DE]-L-A-D-L-T-V-[TA]-N-D-N-D-L-S-[CM]-D-[LM]-[SQ]-D-[SN]-[KR]-[DE]-[AE]-F-K-K-T-W-N-P-[KR]-F-T-L-R-S-H-[FY]-D-[GA]-[IV]-R-[SA]-L-[AT]-F-H-[HPR]-[SG]	D-V-F-I-M-D-[TA]-[IV]
zCM VIIb	[QE]-[SA]-[VA]-L-L-T-A-S-E-D-G-T-L-K-L-W-N-L-[NQ]-K-[AT]-[VAGM]-[ATH]-[AS]-[KR]-K-N-A-A-L-D-V-E-P-I-[HY]-[AT]-F-R-A-H-[RS]-G-[PA]-V-L-[AS]-[VL]-[AT]-[MV]-G-[SE]-[NH]-[SG]-[DE]-X-C-Y-S-G-G-[AL]-D-[AG]	[QE]-[SA]-[VA]-L-L-T-A-S-E
zCM VIIc	[CRS]-[IV]-[HR]-[CSG]-W-[KR]-[IM]-P-D-L-[NSH]-[MV]-D-P-Y-D-[GN]-Y	[CRS]-[IV]-[HR]-[CSG]-W-[KR]-[IM]-P
zCM VIII	D-P-[SG]-[VI]-[LE]-S-[HSC]-V-L-[AE]-G-H-[GE]-D-[AS]-V-W-G-L-[AT]-[FY]-S-[PST]-[AFT]-[SH]-[QH]-R-L-A-S-C-S-A-D-G-[ST]-[IV]-R-[iv]-W-[DN]-P-[SQ]-[SGN]-[SD]	D - P - [S G] - [V I] - [L E] - S - [H S C] - V - L
zCM IXa	C-L-[CS]-[TV]-F-X-[MTAK]-[ATE]-[SGR]-[DGE]-H-[DG]-X-P-T-S-V-A-F-[TV]-[SA]-T-[DE]-P-[AN]-[HQ]-X-V-[AV]-[SA]-F-[RD]-[SG]-G-[DE]-T-V-L-Y-[DN]-[ML]-[ET]-[ATV]-[GE]-[SGQ]-[AS]-[LT]-[LT]-[AST]-L-[DE]-[ST]	C-L-[CS]-[TV]-F-X-[MTAK]-[ATE]-[SGR]-[GDE]
zCM IXb	[RQ]-[GT]-[SNK]-[SD]-G-[PS]-[ATE]-[LQ]-I-N-[QR]-V-[VA]-S-H-P-[SN]-[QE]-P-[LV]-[ST]-I-T-A-H-[DE]-[DN]-R-[GT]-I-R-F-L-D-[NS]-[RN]-T-G-K-X-V-H-S-M	[RQ]-[GT]-[SNK]-[SD]-G-[PS]-[ATE]-[LQ]-I

However, certain members such as striatin in *P. abelii* and zinedin in *O. cuniculus* lacked these domains, either due to variation in the N-terminus region or due to incomplete protein sequence, creating gaps in sequence alignment analysis and resulting in no identification conserved motifs in these members. Although the significance is unclear as yet, SG2NA has two of these domains—sgCM I and sgCM II—well conserved throughout all metazoan homologs.

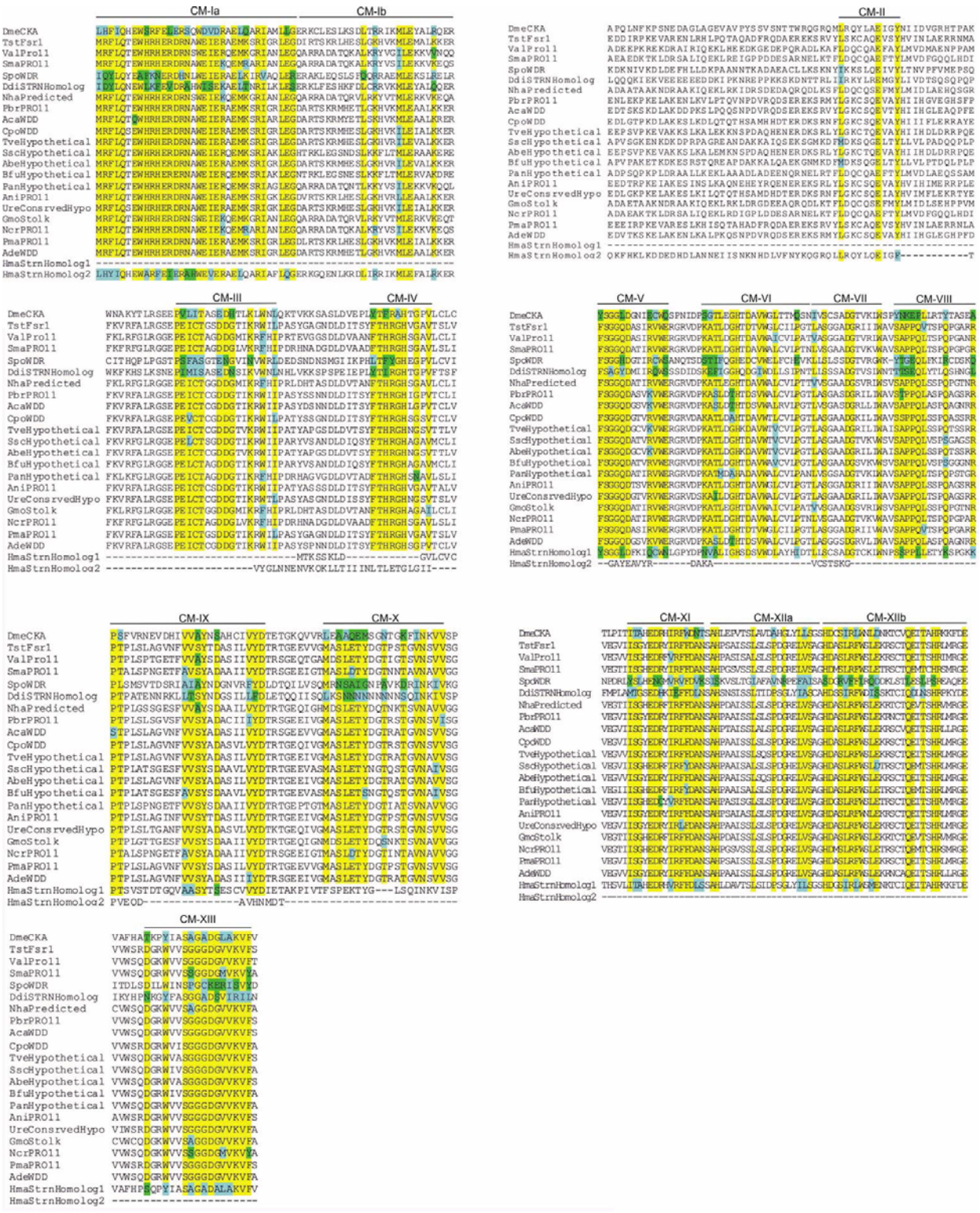


Figure 6. Motif conservation in striatin family members in lower eukaryotes. Sequences of striatin family of proteins from lower eukaryotes were aligned by MAFFT program, and gaps were removed. Thirteen conserved motifs were identified. CM I and XII were divided into two submotifs for further analysis. Among two homologs in hydra, HmaStrn-I does not have CM I–IV and CM XI, whereas HmaStrn-II does not have CM III–XIII.

Table 8. Conserved motifs in the lower eukaryotes (leCM)

Conserved motifs	Pattern	Minimum region
leCM Ia	[MLI]-[HRQD]-[FY]-[LI]-Q-X-[QE]-[WA]-[HSFL]-[RK]-[HNF]-E-[VRL]-[DE]-[RH]-[NAS]-X-W-[DEI]-[VIS]-[DE]-[RK]-X-E-[ML]-[KQRT]-X-R-[IV]-[AGL]-X-L-[LE]-[SG]	[MLI]-[HRQD]-[FY]-[LI]-Q-X-[QE]-[WA]-[HSFL]-[RK]
leCM Ib	[LQ]-X-[KR]-[RHYF]-[IVAL]-X-[MI]-L-E-X-X-[LAV]-[KVRQ]-[QDKNE]-[QELR]-[RQSALT]	[LQ]-X-[KR]-[RHYF]-[IVAL]-X-[MI]-L-E-X-X-[LAV]-[KVRQ]-[QDKNE]-[QELR]
leCM II	[LIM]-[RSDKG]-[QK]-[CYS]-[LAQSTG]-[QARSG]-E-[IVFML]-[YF]-[NHLI]	Nonspecific
leCM III	P-[VESI]-[LIFMV]-[ICAT]-[ST]-[AGCS]-[SGT]-[DE]-[DN]-[HGN]-[TMLVS]-[LIV]-K-[RVL]-[WF]-[NIH]-[LI]	P-[VESI]-[LIFMV]-[ICAT]-[ST]-[AGCS]-[SGT]-[DE]-[DN]-[HGN]-[TMLVS]-[LIV]-K-[RVL]
leCM IV	[YFL]-T-[FHI]-R-[AG]-H-X-[GN]-[PAS]-[VI]	[YFL]-T-[FHI]-R-[AG]-H-X-[GN]-[PAS]-[VI]
leCM V	[YF]-S-[GA]-G-[QHYL]-D-[GAMTF]-[NSTCK]-[IV]-[ERK]-[VCQ]-W-[ENSQ]	[YF]-S-[GA]-G-[QHYL]-D-[GAMTF]-[NSTCK]-[IV]-[ERK]
leCM VI	[SKN]-[GATEV]-[STIFA]-[LFIM]-[DEIGQ]-[GTA]-H-[TEQK]-D-X-[IV]-W-[GAEDT]-[LIV]-[TCFLRA]-[TIVSCY]-[MLHI]-[PQEI]-X-[NTK]	[SKN]-[GATEV]-[STIFA]-[LFIM]-[DEIGQ]-[GTA]-H-[TEQK]-D-X-[IV]-W
leCM VII	[ILV]-[VAL]-S-[CGLS]-[SGA]-[SA]-D-G-[IRST]-[IVC]-[KIRSL]-[LVGI]-W-[SAKN]	[ILV]-[VAL]-S-[CGLS]-[SGA]-[SA]-D-G-[IRST]-[IVC]-[KIRSL]-[LVGI]-W-[SAKN]
leCM VIII	[YST]-[NATS]-[KPGS]-[EP]-[PQL]-X-[RSKT]-[TPILY]-[YQRSK]-X-[AGDHSP]-[SAPNTG]-X-[ARQLK]	Nonspecific
leCM IX	[IFL]-[VAL]-[VITA]-[AS]-Y-[NTSA]-[DS]-[AGE]-[HSANC]-[CIV]-[ILRV]-[VFLI]-[YF]-D-[TLI]-[ERD]-[ST]-[GQA]-[KEIQ]-[QELIP]-[VTI]-[VGAS]-X-[LMF]	[IFL]-[VAL]-[VITA]-[AS]-Y-[NTSA]-[DS]-[AGE]-[HSANC]-[CIV]- [ILRV]
leCM X	[ASN]-[ALSN]-[DEAQN]-[TEIN]-[MYGNS]-[DSN]-[GPNQ]-[NTAS]-X-[GSNKA]-[KTDS]-[GAQRF]-[IV]-N-[KSA]-[IV]-[IV]	[ASN]-[ALSN]-[DEAQN]-[TEIN]-[MYGNS]-[DSN]-[GPNQ]-[NTAS]-X-[GSNKA]-[KTDS]-[GAQRF]-[IV]-N-[KSA]
leCM XI	[IYM]-[ST]-[AGL]-[HY]-E-[DN]-[RGHC]-X-[IV]-[RE]-[FVL]-[FWY]-D-[NAVL]-[NTKS]	[IYM]-[ST]-[AGL]-[HY]-E-[DN]-[RGHC]-X-[IV]-[RE] -[FVL]
leCM XIIa	[AI]-[HS]-[PKSL]-X-[PAS]-[IV]-[STA]-[SGA]-[IL]-[STA]-[VLFIQ]-[SDA]-[APV]-[HDNS]-[GR]-[LRP]-[EY]-[LFI]-[VLA]-[SI]-[GASC]-[SGA]	[AI]-[HS]-[PKSL]-X-[PAS]-[IV]-[STA]-[SGA]-[IL]-[STA]-[VLFIQ]-[SDA]-[APV]-[HDNS]-[GR]-[LRP]-[EY]-[LFI]- -[VLA]
leCM XIIb	[HS]-D-[CAMGS]-[SR]-[ILV]-[RF]-[LF]-[WL]-[NSRD]-[LQIM]-[DES]-[NKTD]-[KR]-[STLN]-[CS]-[TVAI]-[QL]-[DE]-[IMSLV]-[TLN]-[ASCP]-[HS]-R-X-[KMAL]-[RFQY]-[GDE]-E	[HS]-D-[CAMGS]-[SR]-[ILV]-[RF]-[LF]-[WL]-[NSRD]-[LQIM]-[DES]-[NKTD]-[KR]-[STLN]-[CS]-[TVAI]-[QL]-[DE]-[IMSLV]-[TLN]-[ASCP]-[HS]-R
leCM XIII	[TDNS]-X-X-[WY]-X-X-S-[AGSP]-G-[AGC]-[DK]-[GESA]-[VRML]-[VAI]-[KRS]-[IV]-[FYI]	[TDNS]-X-X-[WY]-X-X-S-[AGSP]-G-[AGC]-[DK]-[GESA]-[VRML]-[VAI]-[KRS]

Table 9. Conserved motifs in the C-terminal domain of Archaea and bacteria (prokaryotes)

Conserved motifs present in Archaea and bacteria	Pattern
pCM I	[SNVI]-X-[SADN]-X-[MSATQ]-[IFVW]-[SATH]-[DRPS]-[KYDN]-[VSG]-X-X-X-[IFLAV]-[LNST]-[DEGA]-[DHGS]
pCM II	[TSA]-[QGSAL]-[WS]-X-[DE]-X-[DTSEA]-X-X-X-[IW]-[PNDS]
pCM III	[HPQ]-X-X-X-[IV]-X-X-[VLAC]-[VSAEI]-[IFVWA]-[TSH]-[STPL]-[ND]-G-X-X-X-[IVA]-[ST]-[ACG]-[SG]-X-D-X-[TFSL]-X-[RKA]-[IVL]-W-[DENS]
pCM IV	H-X-X-X-[VIPL]-X-X-[IVLC]-[ASVTE]-[VIFW]-[STH]-[LPA]-D-[GE]-[KQRH]-X-[ALVI]-[AVLI]-X-[CGAVW]-S

Although striatin-like proteins were identified in both nonmetazoans and prokaryotes, further analysis showed that the homology was restricted to the WD-40 domains. The conserved motifs identified in the metazoan proteins were completely absent in their prokaryotic and nonmetazoan counterparts. Taken together, striatin family members appear to have evolved from their prokaryotic counterparts by acquiring domains exclusive for metazoans. Such derivation is further substantiated by the observation that conserved motifs such as those present in the nonmetazoan and prokaryotic WD-40 counterparts were identifiable in SG2NA but not in striatin and zinedin in ScanProsite search. WD-40 domain–containing proteins are widely present in eukaryote and prokaryotes. Although the function assigned for this domain is same across, its evolution when considered in particular may be altogether a different study. However, the conserved motif analysis of prokaryotic striatin homologs, at the C-terminus region, reveals SG2NA acquiring a prototype of bacterial WD-40 domain. The conserved motif sgCM VIIId encompassing the WD repeat domain of SG2NA was also identified in two prokaryotic hits in ScanProsite analysis. With these results, it could be hypothesized that SG2NA might be the earliest among the striatin family to evolve; and at some point in evolution, there were perhaps lateral transfers of sequences between SG2NA and prokaryotic WD repeat proteins

5. METHODS

5.1. Sequence analysis of striatin, SG2NA, and zinedin

Amino acid sequences of human and mouse striatin, SG2NA, zinedin, and their potential variants were retrieved from Ensemble and NCBI databases. To identify homologues of striatin in higher eukaryotes, the full-length human striatin sequence was used as a query and BLASTP analysis was performed against NCBI database of non-redundant protein (nr) sequences. Sequences with more than 40% homology were selected and further validated by using as a query sequence in BLASTP analysis against the NCBI database of human nr sequences. Sequences homologous to SG2NA and zinedin in higher eukaryotes were also identified by the same method. To identify homologs of striatin, SG2NA, and zinedin in lower

eukaryotes, BLASTP analysis was performed by using full-length human striatin sequence; only those with maximum scores were selected. The list of metazoan organisms included in this study is given in Table 7. Lists of nonmetazoan and prokaryotic organisms are given in Tables 8 and 9, respectively.

5.2. Construction of phylogenetic tree

Jalview version 2.6.1 was used for multiple sequence analysis as well as for removing gaps from aligned sequences (32). Multiple sequence alignment was performed by using MAFFT (33). To construct the phylogenetic tree of the variants of striatin family members, sequences from human, mouse, and rat (data not shown) were aligned by MAFFT multiple alignment program; gaps were removed; and the tree was created by Neighbour Joining method (33).

Domain-wise alignment method was used for constructing the phylogenetic tree for the members of the striatin family from higher and lower eukaryotes. The N-terminal and the C-terminal domains of all the family members were separately aligned by using MAFFT; gaps were removed; and the phylogenetic tree was created by Neighbour Joining method.

5.3. Identification of conserved motifs

The conserved motifs were identified manually after aligning the sequences by using MAFFT. The presence of those motifs was then validated by Gblock alignment (34). The sequences were aligned by using MUSCLE (multiple sequences alignment), and the aligned sequences were analyzed by using Gblock (34-35). To determine the specificity of each of the conserved motifs, ScanProsite was used with match mode of greedy, overlaps, and no includes. To identify the core conserved motifs, each of the motifs were restricted by deleting amino acid sequences from the C-terminal ends and analyzing against UniProtKB/Swiss-Prot (including splice variants), UniProtKB/TrEMBL, and PDB using ScanProsite (36).

6. ACKNOWLEDGMENTS

Goutam Kumar Tanti and Nandini Singarapu equally contributed to data analysis. S K Goswami and Rohini Muthuswami equally contributed to data interpretation and manuscript writing. We thank Prof. Alok Bhattacharya for critically review of the manuscript. GKT is a recipient of fellowship from CSIR, Govt. of India. This study was done with the financial support from the Department of Science and Technology, Govt. of India (Grant No SR/SO/AS-37/2009) and the Department of Biotechnology, Govt. of India (DBT-BUILDER_JNU).

7. REFERENCES

1. E J Neer, CJ Schmidt, R Nambudripad and TF Smith: The ancient regulatory-protein family of WD-repeat proteins. *Nature* 371(6495), 297-300 (1994)

2. L van der Voorn and HL Ploegh: The WD-40 repeat. *FEBS Lett* 307(2), 131-4 (1992)

3. D Li and R Roberts: WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell Mol Life Sci* 58(14), 2085-97 (2001)

4. C Xu and J Min: Structure and function of WD40 domain proteins. *Protein Cell* 2(3), 202-14 (2011)

5. N Pashkova, L Gakhar, SC Winistorfer, L Yu, S Ramaswamy and RC Piper: WD40 repeat propellers define a ubiquitin-binding domain that regulates turnover of F box proteins. *Mol Cell* 40(3), 433-43 (2010)

6. Z Shen, KM Sathyan, Y Geng, R Zheng, A Chakraborty, B Freeman, F Wang, KV Prasanth and SG Prasanth: A WD-repeat protein stabilizes ORC binding to chromatin. *Mol Cell* 40(1), 99-111 (2010)

7. S Ritterhoff, CM Farah, J Grabitzki, G Lochnit, AV Skurat and ML Schmitz: The WD40-repeat protein Han11 functions as a scaffold protein to control HIPK2 and MEKK1 kinase functions. *EMBO J* 29(22), 3750-61 (2010)

8. Y Yamamoto, S Mochida, N Miyazaki, K Kawai, K Fujikura, T Kurooka, K Iwasaki and T Sakisaka: Tomosyn inhibits synaptotagmin-1-mediated step of Ca2+-dependent neurotransmitter release through its N-terminal WD40 repeats. *J Biol Chem* 285(52), 40943-55 (2010)

9. CJ Zeng, YR Lee and B Liu: The WD40 repeat protein NEDD1 functions in microtubule organization during cell division in Arabidopsis thaliana. *Plant Cell* 21(4), 1129-40 (2009)

10. H Yu: Cdc20: a WD40 activator for a cell cycle degradation machine. *Mol Cell* 27(1), 3-16 (2007)

11. C. U. Stirnimann, E. Petsalaki, R. B. Russell and C. W. Muller: WD40 proteins propel cellular networks. *Trends Biochem Sci* 35(10), 565-74 (2010)

12. M Benoist, S Gaillard and F Castets: The striatin family: a new signaling platform in dendritic spines. *J Physiol Paris* 99(2-3), 146-53 (2006)

13. F Castets, M. Bartoli, JV Barnier, G Baillat, P Salin, A Moqrich, JP Bourgeois, F Denizot, G Rougon, G Calothy and A Monneron: A novel calmodulin-binding protein, belonging to the WD-repeat family, is localized in dendrites of a subset of CNS neurons. *J Cell Biol* 134(4), 1051-62 (1996)

14. Y Muro, EK Chan, G Landberg and EM Tan: A cell-cycle nuclear autoantigen containing WD-40 motifs expressed mainly in S and G2 phase cells. *Biochem Biophys Res Commun* 207(3), 1029-37 (1995)

15. F Castets, T Rakitina, S Gaillard, A Moqrich, MG Mattei and A Monneron: Zinedin, SG2NA, and striatin are calmodulin-binding, WD repeat proteins principally expressed in the brain. *J Biol Chem* 275(26), 19970-7 (2000)

16. S Gaillard, M Bartoli, F Castets and A Monneron: Striatin, a calmodulin-dependent scaffolding protein, directly binds caveolin-1. *FEBS Lett* 508(1), 49-52 (2001)

17. CS Moreno, WS Lane and DC Pallas: A mammalian homolog of yeast MOB1 is both a member and a putative substrate of striatin family-protein phosphatase 2A complexes. *J Biol Chem* 276(26), 24253-60 (2001)

18. G. Baillat, S. Gaillard, F. Castets and A. Monneron: Interactions of phocein with nucleoside-diphosphate kinase, Eps15, and Dynamin I. *J Biol Chem* 277(21), 18961-6 (2002)

19. V Boonyaratanakornkit: Scaffolding proteins mediating membrane-initiated extra-nuclear actions of estrogen receptor. *Steroids* 76(9), 877-84 (2011)

20. I Letunic, T Doerks and P Bork: SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40(Database issue), D302-5 (2012)

21. M Sanghamitra, I Talukder, N Singarapu, KV Sindhu, S Kateriya and SK Goswami: WD-40 repeat protein SG2NA has multiple splice variants with tissue restricted and growth responsive properties. *Gene* 420(1), 48-56 (2008)

22. P Dutta, GK Tanti, S Sharma, SK Goswami, SS Komath, MW Mayo, JW Hockensmith and R Muthuswami: Global Epigenetic Changes Induced by SWI2/SNF2 Inhibitors Characterize Neomycin Resistant Mammalian Cells. *PLoS One* 7 (11), e49822 (2012)

23. HW Chen, MJ Marinissen, SW Oh, X Chen, M Melnick, N Perrimon, JS Gutkind and SX Hou: CKA, a novel multidomain protein, regulates the JUN N-terminal kinase signal transduction pathway in Drosophila. *Mol Cell Biol* 22(6), 1792-803 (2002)

24. N Oswal, NS Sahni, A Bhattacharya, SS Komath and R Muthuswami: Unique motifs identify PIG-A proteins from glycosyltransferases of the GT4 family. *BMC Evol Biol* 8, 168 (2008)

25. AJ Te Velthuis, T Isogai, L Gerrits and CP Bagowski: Insights into the molecular evolution of the PDZ/LIM family and identification of a novel conserved protein motif. *PLoS One* 2(2), e189 (2007)

26. S Gaillard, Y Bailly, M Benoist, T Rakitina, JP Kessler, L Fronzaroli-Molinieres, B Dargent and F Castets: Targeting of proteins of the striatin family to dendritic spines: role of the coiled-coil domain. *Traffic* 7(1), 74-84 (2006)

27. WB Shim, US Sagaram, YE Choi, J So, HH Wilkinson and YW Lee: FSR1 is essential for virulence and female fertility in Fusarium verticillioides and F. graminearum. *Mol Plant Microbe Interact* 19(7), 725-33 (2006)

28. Y Yamamura and WB Shim: The coiled-coil protein-binding motif in Fusarium verticillioides Fsr1 is essential for maize stalk rot virulence. *Microbiology* 154(Pt 6), 1637-45 (2008)

29. CL Wang, WB Shim and BD Shaw: Aspergillus nidulans striatin (StrA) mediates sexual development and localizes to the endoplasmic reticulum. *Fungal Genet Biol* 47(10), 789-99 (2010)

30. MJ Mizianty and L Kurgan: Modular prediction of protein structural classes from sequences of twilight-zone identity with predicting sequences. *BMC Bioinformatics* 10, 414 (2009)

31. M Levy, Q Wang, R Kaspi, MP Parrella and S Abel: Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant J* 43(1), 79-96 (2005)

32. AM Waterhouse, JB Procter, DM Martin, M Clamp and GJ Barton: Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9), 1189-91 (2009)

33. K Katoh, K Kuma, H Toh and T Miyata: MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33(2), 511-8 (2005)

34. J Castresana: Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17(4), 540-52 (2000)

35. RC Edgar: MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113 (2004)

36. Ede Castro, CJ Sigrist, A Gattiker, V Bulliard, PS Langendijk-Genevaux, E Gasteiger, A Bairoch and N Hulo: ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 34 (Web Server issue), W362-5 (2006)

Abbreviations: Strn: Striatin, CM: Conserved Motif, WD: Tryptophan- Aspartic Acid, eCM: conserved motif in eukaryotes, sCM: conserved motif in striatin, sgCM: conserved motif in SG2NA, zCM: Conserved motif in zinedin, leCM: conserved motif in lower eukaryotes, pCM: conserved motif in prokaryotes.

Key Words: Striatin, Phylogenetic analysis, Conserved motif, nervous system, PP2A, Scaffolding protein.

Send correspondence to: Shyamal K Goswami, Lab No 220, School of Life Sciences, Jawaharlal Nehru University, New Delhi-110067, Tel: 91-11-26704115, Fax: 91-11-26742558, E-mail: skgoswami@mail.jnu.ac.in