

Molecular evolution of proteins mediating mitochondrial fission–fusion dynamics

Sansrity Sinha and Narayanan Manoj

Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai, India

Correspondence

N. Manoj, Department of Biotechnology,
 Bhupat and Jyoti Mehta School of
 Biosciences, Indian Institute of Technology
 Madras, Chennai – 600036, India
 Tel: +914422574113
 E-mail: nmanoj@iitm.ac.in

(Received 27 November 2018, revised 2
 March 2019, accepted 7 March 2019,
 available online 22 March 2019)

doi:10.1002/1873-3468.13356

Edited by Miguel De la Rosa

Eukaryotes employ a subset of dynamins to mediate mitochondrial fusion and fission dynamics. Here we report the molecular evolution and diversification of the dynamin-related mitochondrial proteins that drive the fission (Drp1) and the fusion processes (mitofusin and OPA1). We demonstrate that the three paralogs emerged concurrently in an early mitochondriate eukaryotic ancestor. Furthermore, multiple independent duplication events from an ancestral bifunctional fission protein gave rise to specialized fission proteins. The evolutionary history of these proteins is marked by transformations that include independent gain and loss events occurring at the levels of entire genes, specific functional domains, and intronic regions. The domain level variations primarily comprise loss–gain of lineage specific domains that are present in the terminal regions of the sequences.

Keywords: Drp1; dynamin-related proteins; mitochondrial fission and fusion; mitofusin; molecular evolution; OPA1

Mitochondria are dynamic, bilayered membrane-bound, self-replicating, energy-producing organelles of eukaryotic cells derived from the endosymbiosis of purple non-sulfur bacteria by an ancestral host cell [1–6]. Subsequent to the endosymbiotic event, mitochondria have undergone substantial evolution in terms of metabolic capabilities and ultrastructure [7–11]. The Dynamin-related proteins (DRPs) that belong to the Dynamin superfamily are essential to determine the mitochondrial shape and to regulate remodeling by mediating membrane fission–fusion dynamics [12–16]. Mitochondrial fission/division is regulated by the Dynamin-related protein 1, Drp1 [12,13,17–21]. The mechanism of fusion requires the coordinated sequential fusion of the outer membranes (OMM) and the inner membranes (IMM) induced by the Mitofusin (Mfn1/2) proteins and the Optic atrophy 1 protein (OPA1), respectively [17,22–29]. In yeast, OPA1 (Mgm1) is known to maintain mitochondrial ultrastructure and morphology, in addition to regulating fusion [30,31]. In general, these

proteins undergo GTPase dependent conformational changes to mediate mitochondrial membrane remodeling [32–34].

An early comparative analysis of fission and fusion Dynamin superfamily protein (DSP) sequences suggested a prokaryotic origin of a common ancestor of proteins regulating vesicle and organellar dynamics in eukaryotes. However, the mode of diversification of the DSPs across eukaryotic lineages after the gene duplication event in the last eukaryotic common ancestor (LECA) was unclear due to the limited dataset used [20]. A subsequent large scale comparative study of functionally diverse DSPs proposed multiple duplication events that led to the divergence of the present day DRPs [35]. Purkanti and Thattai proposed that an ancestral bifunctional mitochondrial fission dynamin (hereafter referred to as Drp/Dyn) that existed in the LECA, also mediated vesicle scission. The bifunctional dynamin duplicated into specialized vesicle scission protein (the classical dynamin, Dyn1) and the

Abbreviations

DSP, dynamin superfamily proteins; DRP, dynamin-related proteins.

mitochondrial fission protein independently in the metazoa, green algae, and the alveolates. Primarily, patterns of conservation and divergence of fission DSPs at the level of protein sequences represented as short conserved segments were used to deduce their evolutionary relationships across eukaryotes. However, the study excluded the highly divergent mitochondrial fusion DSPs and bacterial-dynamin-like proteins (BDLPs).

Here we present a comprehensive evolutionary analysis of the DRPs involved in mitochondrial morphodynamics by including the fission and fusion proteins, the prokaryotic BDLPs, and the classical vesicle scission dynamins. In our method, we have incorporated data from conservation patterns of gene splice sites in addition to data from protein sequences. We show that the LECA likely encoded the bifunctional fission protein (Drp/Dyn) and both the fusion proteins (OPA1 and Mfn1). Our results uncover distinct sequence level diversification that may be correlated with the evolution of mitochondrial dynamics and ultrastructure in specific eukaryotic lineages.

Materials and methods

Data retrieval, determination of domain architecture, and multiple sequence alignment

We identified putative full-length orthologs and paralogs of DSPs from all eukaryotic supergroups through several rounds of exhaustive searches using annotated DSP proteins as queries. These include members from Animalia, Fungi, Angiospermae, and basal eukaryotes (apuzoan, Amoebozoa, Choanoflagellata, chromalveolates, excavates, and rhodophytes). The NCBI database was used as a source of sequences. Homologs were identified based on sequence similarity (NCBI BLAST tools, namely, BLASTp, PSI-BLAST, and tBLASTn), domain architecture, and subcellular localization (Text S1). The multiple sequence alignment (MSA) was generated using MAFFT v7.245. (G-INS-1 strategy). BIOEDIT7.25 software was used to edit the alignment to include only the GTPase region in the alignment. Sequences retrieved in this study and the corresponding subcellular localization are listed in Table S1. File S1 contains the full protein alignment of all the sequences in the dataset. Esript3 (<http://esript.ibcp.fr>) was used to display the multiple sequence alignment [36].

Phylogenetic analyses and ancestral sequence reconstruction

Phylogenetic analysis was performed for the full-length protein alignment and the region encompassing the GTPase

domain alone. The selection for the best model [LG model, based on AIC (Akaike information criterion)] and gamma parameter was done using MEGA6.0 [37,38]. Branch position was optimized with a nearest neighbor interchange (NNI) algorithm. The phylogenetic trees were generated based on the maximum likelihood (ML) method and Bayesian strategies using RAXML and MR.BAYES v.3.2.2 programs. In Bayesian analysis, optimal tree topology and posterior probability (PP) values for the nodes, with 100 000 000 Markov Chain Monte Carlo generations using stop value of 0.01 and the burn-in value were determined graphically by removing trees before the plateau. The trees were visualized using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) and iTOL (<http://itol.embl.de/>) [39]. The methodology used is shown in Fig. S1. FastML server [40,41] was used to determine the ancestral sequence of LECA using default settings and LG model, using both marginal and joint reconstruction strategies. The ancestral sequences were thus analyzed to define the domain architecture of the sequence in the LECA.

Gene structure analysis

The exon–intron organization of all genes was determined using the GENE STRUCTURE DISPLAY SERVER (GSDS) software (<http://gsds.cbi.pku.edu.cn/>) [42] and splice sites, intronic phases, and amino acids corresponding to the sequence of each sequence were marked in the alignment file obtained in the multiple sequence alignment obtained earlier (File S1).

Functional divergence tests

Functional divergences (type I and type II) between clusters were tested using DIVERGE 2.0 software [43]. Pairwise coefficients ($\theta_{ij} \pm SE$) and likelihood ratio statistics (LRT) using metazoan Mfn sequences (Mfn1/Mfn2 and Mfn from tetrapods and protostomes) and OPA1 sequences from fungi and metazoa, respectively, were computed.

Results

Identification and origin of mitochondrial fission and fusion DSPs

In keeping with the available evidence for the origin and evolution of the DSPs, a recent classification scheme refers to the DRPs (namely, Drp/Dyn, Drp1, Mfn, and OPA1) as ‘ancient’ dynamins, whereas the classical metazoan vesicle scission dynamins (Dyn1) are termed ‘modern’ dynamins’ [44]. In our study, we follow the above scheme for the definition of the DSP orthologs and paralogs. The orthologs of metazoan vesicle scission dynamins, ancient bifunctional

dynamins and specialized mitochondrial fission dynamins are referred to as Dyn1, Drp/Dyn, and Drp1, respectively. The OMM and IMM fusion mitofusin orthologs are referred to as Mfn or Mfn1/2 and OPA1, respectively.

Orthologs of fission DRPs were identified in both unikont (apusozoans, oomycetes, amoebozoans) and bikont (rhodophytes, chromalveolates, excavates) lineages suggesting that the gene first emerged in the ancestor of these lineages and probably in the LECA. In the unikonts, specialized mitochondrial fission Drp1 orthologs were present only in opisthokonts, whereas unambiguous bifunctional Drp/Dyn orthologs were identified in the Amoebozoa [35,45–47]. The corresponding orthologs in choanoflagellates and apusozoans (represented as Drp1(?) in Table S1) are predicted to be specialized mitochondrial fission dynamins. The single uncharacterized Drp1 ortholog identified in oomycetes may indeed be a specialized Drp1, analogous to that in fungi. Among the bikonts, the specialized fission Drp1 was identified only in the Angiospermeae, whereas all other lineages possessed the bifunctional Drp/Dyn ortholog. The Drp ortholog in the unicellular rhodophyte red algae *Cyanidioschyzon merolae* has been shown to mediate both vesicle and mitochondrial division [48–50]. A similar bifunctional role can be predicted for the corresponding orthologs identified here from other rhodophytes. There is partial, although not conclusive experimental support for predicting a bifunctional role for the orthologs in excavates and chromalveolates. For instance, the sole Drp ortholog in excavate *Trypanosoma brucei* is reported to mediate both vesicle scission and mitochondrial fission, whereas in the mitosome possessing *Giardia lamblia*, the corresponding ortholog was found to colocalize with clathrin-coated vesicles [48].

Unambiguous OPA1 orthologs were identified in the Obazoa and were absent in the Amoebozoa. In bikonts, only the multicellular rhodophyte red algae (*C. crispus*) contained an ortholog of OPA1 in addition to Drp1, whereas all other lineages, namely, Angiospermeae, unicellular rhodophytes, Chromalveolata, and Excavata lacked the OPA1 ortholog. Mitofusin orthologs were identified in the unikonts (opisthokonts) and the bikont (Plantae, including Angiospermeae and rhodophyte *G. sulphuraria*) lineages (Table S1). In the opisthokonts, several lineages contained a single ortholog for Mfn. Interestingly, two homologs (Mfn1 and Mfn2) were identified in the Sarcopterygii (coelacanth and tetrapods). This suggests that a gene duplication event of Mfn2 leading to the emergence of an in-paralog, Mfn1, occurred in an

ancestor of the Sarcopterygii after its divergence from the actinopterygian lineage. In summary, it is evident that the three paralogs that are central to mitochondrial fission and fusion apparatus are present across unikont and bikont lineages and therefore allows for a hypothesis that the LECA most likely contained a set of both fission and fusion DSPs.

Homology of mitochondrial fusion and fission DSPs

Domain addition, deletion, and rearrangements leading to functionally divergent paralogs of the ancestral sequence are commonly observed in the molecular evolution of multidomain proteins. Besides this, conservation of gene structures is also an excellent marker of the evolutionary relationships that characterize orthologous and paralogous genes, subsequent to duplication [51–54]. Hence, a combined analysis of the conservation patterns of protein domain architectures and gene structures were performed for all identified DSPs.

Conservation and divergence of domains

Comparative analyses of domain architectures of the prokaryotic BDLPs and eukaryotic DSP sequences suggests that a conserved five domain architecture, namely, a N-terminal heptad repeat coiled-coil domain (CC0/HRN), the GTPase domain (GTPase), the helical middle domain (MD), the transmembrane (TM) domain, and the helical C-terminal GTPase effector domain (GED) can be considered to constitute the minimal scaffold of an ancestral eukaryotic DSP protein (referred to as the signature DSP domain architecture). Furthermore, all DSPs contain a helical region (NH) connecting the CC0/HRN and the GTPase. The NH region is of varying lengths across the DSPs (Fig. 1). The MD and GED are known to fold back onto each other forming helical bundle regions (HB1/2), also known as BSE/neck and trunk/stalk, respectively. The neck and stalk regions mediate the open and closed conformations of the DSPs upon GTP binding. The two helical bundles are connected by flexible loops that constitute the hinge region. The MD and GED contain two helical segments each (MD1/2, GED1/2), connected by loops. The MD1 and GED2 helices are part of the HB1/BSE/neck structural motif, while the MD2 and GED1 are part of the HB2/trunk/stalk assembly. In the structures of BDLP, Mfn1, Drp1, and Dyn1, the NH region folds back onto the BSE/neck and appears to be a conserved structural feature across DSPs. In the BDLP and Mfn, the NH constitutes two helices, whereas in the Drp1 and

Dyn1, the NH region is much shorter with a single helix. In general, the GTPase and the membrane attachment regions occur on either tip of the neck and stalk regions, respectively (Fig. S2).

The classical vesicle scission Dyn1 orthologs contain the signature DSP domain architecture lacking the CC0/HRN domain. Furthermore, a Pleckstrin homology domain (PH) substitutes for the TM domain and a Dyn1-specific C-terminal proline-rich domain (PRD) is added at the C-terminal end. The PH domain in Dyn1 serves as a site for membrane attachment (Figs 1 and S5), while the PRD binds effector proteins

containing SH3 domain, thereby mediating vesicle scission and endocytosis [55,56].

All fission orthologs (Drp/Dyn and Drp1) contain a variant of the signature DSP domain architecture lacking the CC0/HRN domain. Additionally, variable domain (VD) substitutes for the TM which provide a site for attachment of adaptor proteins to mediate mitochondrial fission [21,57]. All sequences were also examined for signal sequence motifs that correspond to organellar localization patterns. The results obtained here are consistent with previous literature supporting the cytosolic localization of the fission

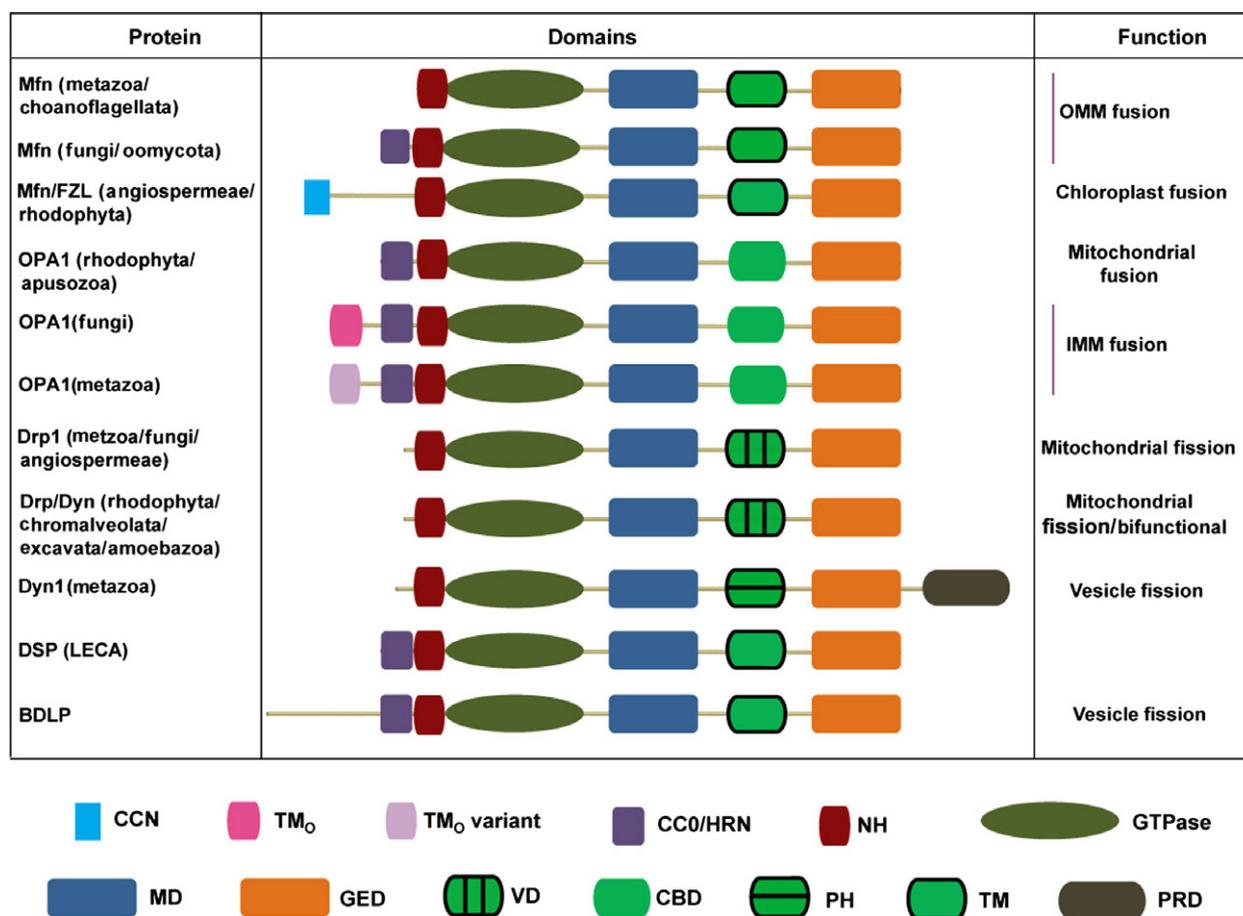


Fig. 1. Domain architecture of Dynamin superfamily proteins. Representation of the domains shown in the schematic here roughly represent the full-length alignment of the amino acid sequences of DSPs and are limited to the defined domains. The left column corresponds to the DSP annotation. Drp1, Mfn, OPA1, Dyn1, and BDLP refer to Dynamin-related protein 1, Optic atrophy1, Mitofusin, vesicle fission dynamins, and bacterial-dynamin-like proteins. The annotations Mfn/FZL refer to Mfn orthologs identified from Plantae lineage and Drp/Dyn refers to the bifunctional (vesicle scission and mitochondrial fission) dynamins identified from specific lineages of unikonts and bikonts. The predicted domain architecture of the ancestral DSP in the LECA was inferred from domain architecture features of DSPs from extant eukaryotic lineages, prokaryotic BDLPs, and the results from ASR (Ancestral sequence reconstruction) analyses. The index for the markers used for representing domains is shown at the bottom of the figure. The highly divergent interdomain regions between CC0/HRN, GTPase, MD, TM/PH/VD/CBD, and GED represented by brown lines are not to scale and do not provide any information regarding the sequence conservation. The column on the right corresponds to the cellular function. The BDLP domain arrangement represents the canonical domain architecture of a prokaryotic DSP.

DSPs and mitochondrial localization of fusion DSPs (Table S1).

Among the mitofusins, the fungal orthologs contain the signature DSP domains, whereas orthologs from other lineages displayed domain level variations. For example, mitofusins in the seed plants, choanoflagellates, and the Metazoa lack the conserved CC0/HRN domain suggesting a lineage-specific domain loss. All mitofusin orthologs possess the TM domain that is equivalent to the PH/VD domain of the Drp1/Dyn1. Interestingly, the TM is also shared between mitofusins and BDLPs. However, unlike in the BDLPs, the TM domain in opisthokont mitofusins harbors the mitochondrial targeting signal sequence (MTS) that is indispensable for mitofusin-mediated OMM fusion (Table S2) [58,59]. The MTS is absent in seed plant and rhodophyte orthologs. Nevertheless, the N-terminal chloroplast targeting signal sequence (CTS) identified in these is consistent with their reported functional role in the regulation of thylakoid networks in the plant chloroplast [20,60]. Hence, the mitofusin ortholog in seed plants and rhodophytes is referred to as Mfn/FZL. Moreover, Mfn/FZL contains a unique nonhomologous N-terminal extended region of about 140 residues that most likely form a coiled-coiled like domain (named here as CCN) with a 60 residue stretch comprising the two heptad repeats. This is suggestive of a lineage-specific gain of the domain in the most recent common ancestor of the Angiospermeae and Rhodophyta.

The OPA1 orthologs also contain the signature domain architecture with a Lys-rich homologous cardiolipin-binding domain (CBD) substituting the PH/VD/TM domain of Dyn1/Drp1/Mfn paralogs (Fig. 1). The CBD in OPA1 has been shown to bind to cardiolipin and mediate inner mitochondrial membrane fusion [30,61–65] (Fig. 1). In the opisthokonts, the OPA1 ortholog contains an N-terminal insertion of a transmembrane domain (TM_O) that is known to be responsible for anchoring OPA1 to the IMM. Upon GTP binding, the region spanning the MTS, and the TM_O that acts as a stop-transfer sequence, is cleaved by IMM peptidases, resulting in the short isoform, s-OPA1. The processing and the balanced formation of the long and short isoforms together are essential for IMM fusion and maintenance of mitochondrial morphology, including the structure of IMM [61,66–69]. Interestingly, in the yeast OPA1 ortholog (Mgm1), the hydrophobicity of TM_O has been shown to determine the processing efficiency by the peptidase Pcp1. Furthermore, changing the hydrophobicity of this segment altered the ratio of the isoforms leading to mitochondrial fragmentation. In stark contrast, the rat

OPA1 processing by the peptidase m-AAA was not modulated by increment of TM_O hydrophobicity, suggesting that the mechanism of processing of l-OPA1 in mammals is distinct from that of yeast l-Mgm1 [66,70].

We, therefore, examined the hydrophobicity profiles of the TM_O domains in the all metazoan and yeast orthologs to look for distinguishing features between the two groups, if any. It is noteworthy that the estimates of transmembrane (TM) probability for the TM_O domain in metazoan OPA1 orthologs (PP = 0.0825) (referred to as TM_O variant) is significantly less than that of fungal Mgm1 orthologs (PP = 0.1716) (P -value < 0.0205, unpaired t -test). Examination of type II functional divergence estimates reveals sites in the TM_O domain where conserved charged amino acids in the metazoan sequences are substituted by hydrophobic amino acids in fungi (K116Y and K117A in human OPA1). These substitutions contribute significantly to the distinct hydrophobicity profiles of the TM_O between the two groups (File S4, Table S3 and Fig. S5d). We propose that these residues may play a critical role in the distinct processing mechanisms of the long isoform in these lineages. Interestingly, the rhodophyte and apusozoan orthologs lack the TM_O, indicating a likely distinct mechanism of processing and IMM fusion in these lineages.

Conservation of exon–intron structures and intradomain splice site positions

Next, the gene structures of all sequences were examined for conservation of exon–intron structures, splice site positions, and phases within the contiguous sequence regions that encode the domains described above. The interdomain regions were excluded from the analysis because of the alignment uncertainties of these regions. The genes encoding fission DSPs, including the bifunctional Drp/Dyn and the specialized Drp1 or Dyn1 displayed lineage dependent conservation patterns of gene structures. The Dyn1 gene contains a highly conserved pattern of intron numbers and positions across the Metazoa (Fig. 2). It is notable that in the bikont lineage, in general, most bifunctional orthologs are encoded by single exon genes (SEG), whereas orthologs that encode specialized dynamins contain introns. In the unikont lineages, the specialized mitochondrial fission Drp1 orthologs in Fungi are encoded by SEGs. The predicted Drp1 orthologs in the apusozoan *Thecamonas trahens* and oomycete *Aphanomyces invadans* (represented as Drp1(?) in Fig. 5) are encoded by intron containing genes although another oomycete, *Phytophthora parasitica*, contains a Drp1 encoded by a

SEG. Remarkably, the Drp1 in seed plants share significant conservation of splice site positions and exon numbers in the GTPase, the MD, and the GED, with their corresponding orthologs in the deuterostomes. The shared splice site positions strongly suggest that the Drp/Dyn in the common ancestor of the Unikonta and Bikonta and presumably, the LECA, encoded an intron containing Drp/Dyn gene. Consequently, the SEGs, for instance, in the fungi in unikonts or excavates in the bikonts, appear to arise from independent intron loss events in these lineages.

OPA1 in the bikont Rhodophyta is encoded by a SEG. Among the unikonts, the fungal orthologs are encoded by SEGs, whereas the apusozoan, choanoflagellate, and metazoan orthologs are encoded by intron containing genes. Furthermore, the intradomain exons observed in the choanoflagellates and metazoan orthologs share conserved features of exon numbers, intronic phases, and splice site positions (Fig. 2). The OPA1-specific TM_O variant domain identified in the opisthokonts (Metazoa, Choanoflagellata, and Fungi) is encoded by a single exon suggesting that the TM_O domain evolved in the OPA1 by an exon insertion event in an ancestor of the opisthokonts and a subsequent intron loss occurred in the lineage leading to the fungi.

The fungal mitofusins are encoded by SEGs, similar to that observed for the OPA1 and Drp1 paralogs. The uncharacterized mitofusin ortholog identified from unicellular rhodophyte *G. sulphuraria* is encoded by a gene containing one intron. The CTS is invariably present in all seed plants (encoded by a single exon) orthologs. The absence of the MTS and presence of the CTS in seed plant Mfn/FZL orthologs proteins enable their targeting to the chloroplasts [60]. Our search for other homologs of the mitofusin across rhodophytes (three complete genomes) and seed plant genomes (12 complete genomes) failed to pick up any hits. Hence, the identity of the mitofusin protein in these lineages remains unknown. A single exon encodes the N-terminal CTS, CCN, and the G1-box of the GTPase domain in rhodophyte Mfn/FZL. A single exon codes for seed plant N-terminal CTS, whereas the CCN domain is encoded by two additional exons suggesting that the MTS/CTS and CCN regions evolved by an exon insertion event in a common ancestor of these two lineages, followed by independent intron losses in the rhodophytes. The intron–exon structure of genes encoding mitofusins was significantly conserved across choanoflagellates, metazoans, and seed plants in terms of exon numbers and splice site residues and positions (Fig. 2). Thus, the well-conserved Mfn gene structure across orthologs in the Unikonts and the Bikonts

indicate that the intronic gene structure of Mfn represents the ancestral state in the unikonts and the bikonts, with intron losses occurring in the fungal lineage.

Unlike in the Dyn1 and Drp1 genes, OPA1 and Mfn encoding genes displayed significant conservation of paralog-specific phases across the Metazoa, Choanoflagellata, and the seed plants suggesting additional selection pressures on intron properties (Fig. S7). The DSP domains, in general, are encoded by nonoverlapping sets of exons. The only exceptions include a fused exon encoding the MTS, CCN, and the G1-box of the GTPase domain in rhodophyte Mfn/FZL ortholog, a fused exon encoding the G5 box of the GTPase and the adjacent MD in seed plant mitofusins, and a fused exon encoding the C-terminal region of GED and N-terminal region of PRD in the metazoan Dyn1 orthologs (Fig. 2).

Next, we compared the patterns of conservation and divergence within the intradomain splice site positions across the vesicular Dyn1 and the fusion and fission DSPs. At least three shared splice positions were clearly identified in the GTPase domain (G1 and G2 box) and GED (Fig. 3). The first conserved splice site position (Site I) at G1 box (corresponding to residue Gln34 of HsapDrp1) is shared across Mfn and OPA1 orthologs. However, the corresponding site is not conserved with the corresponding intron position in Dyn1 (+12 downstream) which shared the splice site with non-mitochondrial dynamins namely Atlastins (ATLs) and Guanylate-binding proteins (GBPs) (File.S6). A similarly shared splice site is also present at the exon encoding the G2 box (Site II) and in the GED (Fig. 3). The positions of other splice sites shared across all DSPs including the Dyn1 are shown in Fig. S6. Together, the patterns of shared intradomain splice site positions across the paralogous Mfn, Drp1, and OPA1 sequences strongly suggest that the three paralogs inherited these sites from an intron containing common ancestor, before the duplication event.

Phylogeny of fission and fusion DSPs

Next, evolutionary relationships of the four paralogous groups were studied using phylogenetic tree inference using two independent alignments of the protein sequences. The first included the full-length sequences (Fig. 4 and File S5), while the second included the region encompassing only the GTPase domain (Fig. S3A). The BDLPs were used as an out-group. Interestingly, both trees had similar topologies suggesting that the sequence divergence within the GTPase domain alone is consistent with the

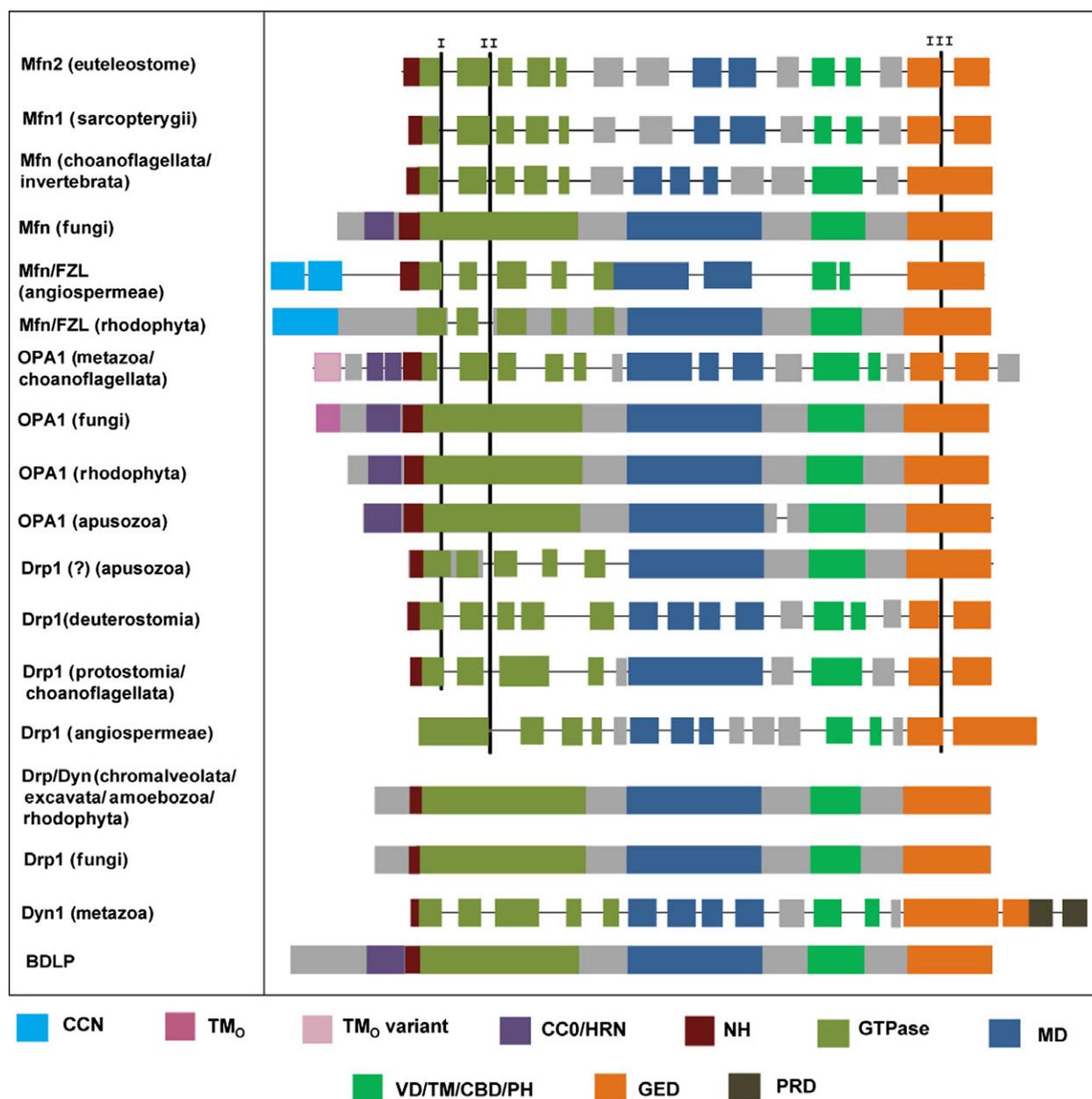


Fig. 2. Gene structures of representative orthologs and paralogs. The left column corresponds to the DSP annotation. Drp1, Mfn, OPA1, Dyn1, and BDLP refer to Dynamin-related protein 1, Optic atrophy1, Mitofusin, vesicle fission dynamins, and bacterial-dynammin-like proteins. The annotations Mfn/FZL refer to Mfn orthologs identified from Plantae lineage and Drp/Dyn refers to the bifunctional (vesicle scission and mitochondrial fission) dynamins identified from specific lineages of unikonts and bikonts. Manually generated schematic representation of gene structures are shown here. Exons corresponding to each domain are represented with different colors as defined in the index shown below the figure. The positioning of domains roughly represents the full-length alignment of the regions coding for the domains. The frequencies of occurrence of colored boxes for a specific domain correspond to the number of exons present in the gene encoding the domain. The domains corresponding to the membrane attachment region, namely the CBD/PH/TM/VD is represented by dark green colored boxes. The BDLP domain schematic shown here represents core domain architecture conserved across most paralogous BDLPs. The black lines represent the intronic regions and the gray boxes correspond to exons in the interdomain regions. The presence of highly divergent interdomain regions limit the sequence-based alignment to well-defined domains. The intronic regions and exons corresponding to interdomain regions are not to scale and do not provide information regarding conservation patterns in these. The shared splice sites named as sites I, II, and III are marked with vertical black lines corresponding to their positions in well-defined domains.

	I	II	III
	GTPase (G1 Box)	GTPase (G2 Box)	GED
HsapMfn1	73-RHMKKVAF <u>F</u> <u>G</u> RTSSGKSSVINAMLWDKVL	131-DEKKS <u>V</u> <u>K</u> T-VNQLAHALHM-DK	706-EKIQNNS <u>K</u> <u>L</u> LRNKAVQLENE
HsapMfn2	94-RNMKKVAF <u>F</u> <u>G</u> RTSNGKSNVINAMLWDKVL	151-EEKRS <u>A</u> <u>K</u> T-VNQLAHALHQ-DK	725-DSLQSK <u>A</u> <u>K</u> <u>L</u> LRNKAGWLDSE
Xlmfn1	75-RNMKKVAF <u>F</u> <u>G</u> RTSSGKSSVINAMLWDKVL	132-EEKKS <u>V</u> <u>K</u> T-VNQLAHALHM-DK	738-EQIQSN <u>S</u> <u>K</u> <u>A</u> LRNRASRIEGE
Xlmfn2	95-RHMKKVAF <u>F</u> <u>G</u> RTSNGKSNVINAMLWDKVL	152-EEKRS <u>V</u> <u>K</u> T-VNQLAHALHQ-DD	705-EQIQSN <u>S</u> <u>K</u> <u>V</u> LRNRASRIEGE
SkMfn	74-RDHMKVAF <u>F</u> <u>G</u> RTSSGKSTVINAMLKDKVL	136-NERRNVRS-VS Q LAHALSN-EK	610-ERIQND <u>A</u> <u>K</u> <u>L</u> LRNKAGWLENE
GmFzo	357-DEPFLV <u>I</u> <u>I</u> <u>V</u> GEFNSSGKSTVINALLGERYL	408-QQ Q CERHP-----DG	888-EIQEELS <u>N</u> VEKKLRTLQIDI
AthFzo	347-DEPFM <u>V</u> <u>I</u> <u>I</u> <u>V</u> GEFNSSGKSTVINALLGKRYL	399-QQ R CQTHP-----DG	913-GIQKELSDIRSKLQLLQVDI
HsapOpa1	286-DHLP <u>R</u> <u>V</u> VVVGDQSSGKSSVLEALVGRDFL	348-FDLT <u>K</u> <u>E</u> D-LAALRHEIEL-RM	930-GKRVQLAED <u>L</u> <u>K</u> <u>K</u> VREIQEKL
DrOpa1	364-DHLP <u>R</u> <u>V</u> VVVGDQSSGKSSVLEALVGRDFL	427-FDLG <u>K</u> <u>E</u> D-LAALRHEIEL-RM	1008-GRRVQLAED <u>L</u> <u>K</u> <u>K</u> VREIQEKL
CgOpa1	305-DNLP <u>R</u> <u>V</u> VVVGDQSSGKTSVLEMIVQARIF	368-FDLT <u>K</u> <u>E</u> S-D-LAALRKEVEL-RM	954-GPRVQLAEE <u>L</u> <u>K</u> <u>K</u> VQRVQIQEKL
HsapDrp1	23-IQLPQIVV <u>V</u> <u>G</u> TQSSGKSSVLEALVGRDFL	96-TKN <u>K</u> <u>L</u> YTD-FDEIRQEIEN-ET-	706-AQRRKEAAD <u>M</u> <u>L</u> <u>K</u> ALQGASQII
DrDrp1	23-IQLPQIVV <u>V</u> <u>G</u> TQSSGKSSVLEALVGRDFL	95-TKN <u>K</u> <u>L</u> YTD-FDEIRQEIEN-ET-	685-AQRREAAD <u>M</u> <u>L</u> <u>K</u> ALQKASQVI
SmDrp1	23-IQLPQIVV <u>L</u> <u>G</u> TQSSGKSSVLENLVGRDFL	95-NK- <u>K</u> IYRD-FDEVRNEIKS-ET-	689-AQRRREASE <u>M</u> <u>L</u> <u>K</u> ALQKASMI
PtDlp1	40-IELPQVAV <u>V</u> <u>G</u> SQSSGKSSVLEALVGRDFL	105-LPGRFYD-FSEIRSEI <u>Q</u> A-ET-	727- <u>A</u> MKRKQTREQLRVLQQAFTL
PvDlp1	45-IDLPQVAV <u>V</u> <u>G</u> SQSSGKSSVLEALVGRDFL	108-LPGRKFHD-FSDIRREI <u>Q</u> A-ET	720- <u>A</u> LKRKRCELLRAYQQAQFDL
SbDlp1	52-IDLPQVAV <u>V</u> <u>G</u> SQSSGKSSVLEALVGRDFL	116-LSGRRFYD-FREIRREI <u>Q</u> A-ET	725- <u>T</u> IKRKQIRENLKVLQQAQYKTL
HsapDyn	29-DLPQIAVVVGGQSSAGKSSVLENF <u>V</u> GRDFL	86-CKSKKFTD-FDEVRQEIEA-ET	711-AQRRDMLRMYHALKEALNII
DrDyn	29-DLPQIAVVVGGQSSAGKSSVLENF <u>V</u> GRDFL	86-CKGRKFVD-FDEVRQEIEA-ET	712-AQRRDEMLRMYHAIKEALSII
IsDyn	24-DLPQIAVVVGGQSSAGKSSVLENF <u>V</u> GRDFL	81-CRGGKFMFD-FDQIRKEIED-ET	715-AQKRDEMLRMYHACKEALRII

Fig. 3. Conserved splice site positions at intradomain exons across mitochondrial fusion and fission DSPs. The figure shows alignment of residues corresponding to the shared intradomain splice site positions in the GTPase domain (G1 and G2 boxes) and GED, within and across DSPs. A 20 residue sliding window (highlighted in gray) was used to assess splice site conservation across the DSPs. The reference residue is highlighted in red font. All splice sites are underlined and marked in bold within each domain. Splice site occurring up to ± 10 positions from the reference residue splice site were considered conserved. The prefixes of the sequence names in the left column, Hsap, Xl, Sk, Dg, Is, Cg, Sm, Gm, Ath, Pt, Pv, Sb correspond to the organisms, namely, *Homo sapiens*, *Xenopus leavis*, *Drosophila melanogaster*, *Ixodes scapularis*, *Crassostrea gigas*, *Schistosoma mansoni*, *Saccoglossus kowalevskii*, *Glycine max*, *Arabidopsis thaliana*, *Populus trichocarpa*, *Phaseolus vulgaris*, and *Sorghum bicolor*, respectively. The numbers preceding the sequence in each column correspond to the position of the initial amino acid of the region of the sequence presented. The shared splice site across sequences in the GTPase domain, and GED are marked with Roman numerals I, II and III.

paralogous relationships between the DSPs and, that the evolution of GTPase domain occurred in parallel with the evolution and diversification of the other regions of the sequences. The tree topology generated from sequence features alone was also congruent with that generated using a structurally informed sequence alignment (Fig. S3B). Overall, the observed pattern of clustering in the phylogenetic tree is consistent with the conservation patterns inferred from gene presence/absence, domain architecture, and intradomain gene structure analyses.

It is evident that the most recent gene duplication event among the three paralogous gene families examined here occurred in an ancestor of the Sarcopterygii leading to the emergence of the in-paralogs, Mfn1 and

Mfn2. This is consistent with the well-supported clusters corresponding to euteleostome Mfn2 and sarcopterygian Mfn1 in the phylogenetic tree. Functionally divergent sites [61 sites (PP > 0.6) displayed type I divergence (θ -I = 0.474400 ± 0.069687 , P -value > 0.05); 26 sites (PP > 1) displayed type II divergence (θ -II = 0.059572 ± 0.037470 , P -value > 0.05)] were identified in all domains across the in-paralog sequences from tetrapods. GTPase domain-mediated dimerization is a shared feature across various GTPases including DSPs, Septins, TRAFECs, Ras-like, and GTPase of immunity-associated proteins (GIMAPs). However, the involvement of G1, G2 boxes, and intervening regions between the G-boxes in the dimerization interface (G-interface) is a specific feature of DSPs alone. Also,

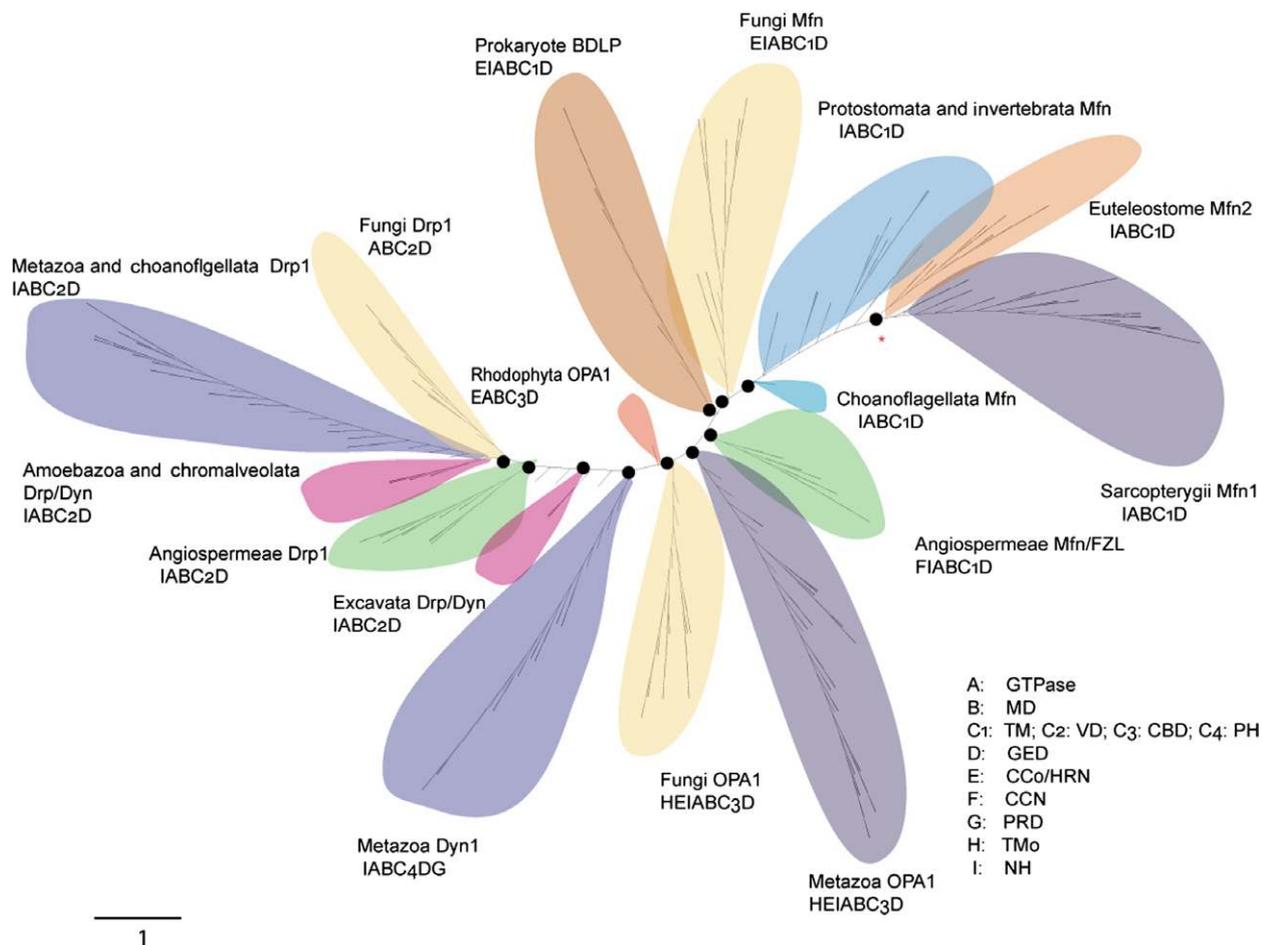


Fig. 4. Phylogenetic relationships of mitochondrial fusion and fission DSPs. The reconciled Bayesian and ML unrooted tree generated from 371 full-length protein sequence alignment of Mfn, Drp1, OPA1, Dyn1, and BDLP sequences. Accession numbers are given in Table S1. The Mfn/FZL from *G. sulphuraria*, OPA1, and Drp/Dyn sequences from apusozoa and oomycete were excluded from the phylogenetic analyses. The branches corresponding to posterior probability <0.8 and bootstrap support <80 are collapsed. The black circles represent the posterior probability and bootstrap values of 1 and 100, respectively, representing common nodes inferred in both Bayesian and ML trees. The image was created using the iTOL web server. The clusters corresponding to each DSP is marked with shaded regions of colors specific for each lineage. The metazoan and fungal DSPs are marked with purple and peach colors, while the seed plant and rhodophyte DSPs are shaded with dark green and red. The choanoflagellate, protostomes, and invertebrate Mfn cluster is marked in light blue, whereas the cluster corresponding to bifunctional Drp/Dyn in basal eukaryotes, namely, the amoebozoans, alveolates, and excavates are marked with pink. The prokaryotic BDLPs represent the outgroup and the cluster corresponding to those is shaded in brown. The distinct clusters of Euteleostome Mfn2 and Sarcopterygii Mfn1 are marked with orange and purple, respectively. The domain architecture corresponding to each lineage is mentioned below the name of the cluster. The index for alphabetical codes used for defining domains in this figure is shown at the bottom right corner of the figure. The scale bar represents 1 substitution per site.

GTPase domain triggered adhesion and cohesion patterns of associated HB1/2 to mediate membrane remodeling are a feature specific to DSPs [71]. Interestingly, our study has identified shared features regarding the G-interface between GIMAPs and Mfn1. Both GTPases share a polar contact between the two molecules on the G-interface (Fig. S4). Interestingly, specific differences in the residues constituting the G-interface or in close proximity to it were identified in Mfn1/2

proteins suggesting possible differences in dimerization rates between the two in-paralogs (Fig. S4, File S4, Table S4, Text S1). Together, these results suggest that these proteins have indeed undergone neofunctionalization.

The tree displays distinct robustly supported clades that correspond to Mfn1/2, OPA1, Drp1, Dyn1, and BDLPs. Among these, the BDLP, Drp1, and the Dyn1 clusters are monophyletic. Additionally, distinct

clusters corresponding to the bifunctional fission DSPs in Amoebozoa, chromalveolates, and excavates were monophyletic. The clustering of sequences identified from basal eukaryotes (including Mfn/FZL ortholog from *G. sulphuraria*, OPA1 ortholog from *C. crispus*, and Drp/Dyn orthologs from amoebozoans, oomycetes, rhodophytes, chromalveolates, and excavates) is consistent with their sequence-feature-based identification and hence confirms the assigned annotations (Fig. S3, File S3 and File S5).

BDLPs and mitofusins share at least four out of five domains of the signature architecture suggesting close structural and functional similarities [72]. In particular, the TM is shared only between the mitofusins and BDLPs. Furthermore, the Mfn and the BDLP clades cluster into a well-supported monophyletic clade (bootstrap value > 90) in the phylogeny. To examine if this relationship is also reflected at the level of the three dimensional structures, crystal structures of BDLP with Mfn1, Drp1, and Dyn1 were compared. A broad comparative analysis of the structure–function relationships between these proteins and in particular of the GTPase domain has been reported previously [32–34]. In our study, the analysis is limited to the MD-GED helical segments of the neck/HB1 since this structural motif is present in the structures of all four paralogs (for instance, the Mfn1 structure is that of a truncation variant lacking the HB2 motif). The overall helical arrangement is significantly conserved (root-mean-square deviation, RMSDs of 0.64–3.1 Å). Furthermore, the hydrophobic core is also well conserved as expected (Fig. S8). However, divergence between these helical regions is present at the level of amino acid residues constituting DSP specific sequence signatures (Fig. S5). Interestingly, BDLP and Mfn1 shared multiple conserved features unique to the two paralogs. These include overall length of the helices, an extended conserved hydrophobic core and the presence of salt bridge interactions that stabilize the helical interface (Fig. S8). The overall HB1 motif has remained largely structurally conserved across DSPs in spite of significant sequence divergence in these regions. We propose that the DSP specific sequence features may affect the nature of packing at the interfaces and modulate the assembly, cohesion and flexibility of these helical regions in their functional contexts.

Discussion

Here we report the molecular evolution of eukaryotic mitochondrial fusion and fission DSPs that relies on combining information from gene presence/absence,

domain conservation patterns, protein sequence phylogeny, and intradomain gene structure features. Given the long time scales in the evolution of these proteins the interdomain regions are highly divergent. Therefore, the power of protein sequence based phylogenetic analysis to infer consistent relationships between the homologs is limited to analyses of regions that encode functional domains. The absence of domain shuffling events in the diverse dataset of DSPs examined here indicates that the signature domain architecture of the DSPs evolved under strong negative selection pressure. We show that diversification after duplication in these paralogs have occurred in following ways: (a) addition and deletion of new domains at the terminal regions (for instance, addition of PRD, TM_O, and CCN in the metazoan Dyn1, Opisthokonta OPA1, and Plantae Mfn/FZL, respectively) (b) sequence level changes translating to functional divergence within the membrane attachment region (TM/VD/PH/CBD variants in Mfn/Drp1/Dyn1/OPA1), and (c) divergent splice site positions and phases within and across paralogs. Furthermore, all three paralogs in the LECA are encoded by intron containing genes and parallel intron loss events appear to have occurred later in specific eukaryotic lineages (Fungi, Rhodophyta, Chromalveolata, and Excavata). Intron loss in specific lineages is in line with the ‘intron-early’ hypothesis and is consistent with similar low intron density, particularly in housekeeping and cell survival proteins in fungi and parasitic eukaryotic lineages [73–77].

Since mitochondria cannot be formed *de novo* the presence of either the specialized Drp1 or the bifunctional Drp/Dyn is indispensable in eukaryotes to allow the symmetric distribution of mitochondria to daughter cells during cell division. We show that the bifunctional Drp/Dyn diversified independently into specialized vesicle scission and mitochondrial/chloroplast fission dynamins in Metazoa and seed plants, whereas it retained its bifunctional role in other lineages that include the Amoebozoa in the unikonts and, in chromalveolates, excavates and rhodophytes in the bikonts. These inferences are congruent with that reported by Purkanti and Thattai [35]. It is noteworthy that the Glaucophytes lack the DRPs responsible for fission and probably recruit the FtsZ_p and FtsZ_m proteins for division of their mitochondria and specialized chloroplast (cyanoplast) (Fig. 5) [35,78]. However, the identity of the paralog mediating mitochondrial fusion in this lineage cannot be established using the available data. The presence of both fusion paralogs (OPA1 and Mfn1), in specific lineages across unikonts and bikonts, albeit, with independent gene losses in amoebozoans, excavates and chromalveolates suggest that

these paralogs emerged concomitantly as a result of a gene duplication in the LECA. The loss of the fusion apparatus in excavates and chromalveolates is consistent with the fact that these lineages contain only a single mitochondria per cell, precluding the need for proteins that mediate mitochondrial fusion [9]. The absence of the fusion apparatus in the Amoebozoa cannot be explained using the data presented here and necessitates further studies.

The bikonts examined here contain both chloroplasts and mitochondria. The identity of the mitochondria-specific mitofusin protein in seed plants remains unknown since the Mfn/FZL ortholog has been shown to target to the chloroplast and not to the

mitochondria [60]. We hypothesize that the insertion event of the Plantae-specific CTS and CCN domains in the Mfn/FZL enabled its neofunctionalization by determining the localization and evolution of its functional role in regulating the morphology of the plant chloroplast thylakoid networks.

The rhodophytes are one of the earliest eukaryotic algae, representing a unique class with unicellular and multicellular members. The rhodophytes present an interesting pattern of loss–gain events of the two fusion paralogs. The sole fusion paralog Mfn/FZL identified in the unicellular rhodophyte *G. sulphuraria* most likely plays a functional role in the formation of thylakoid networks, similar to that observed in seed

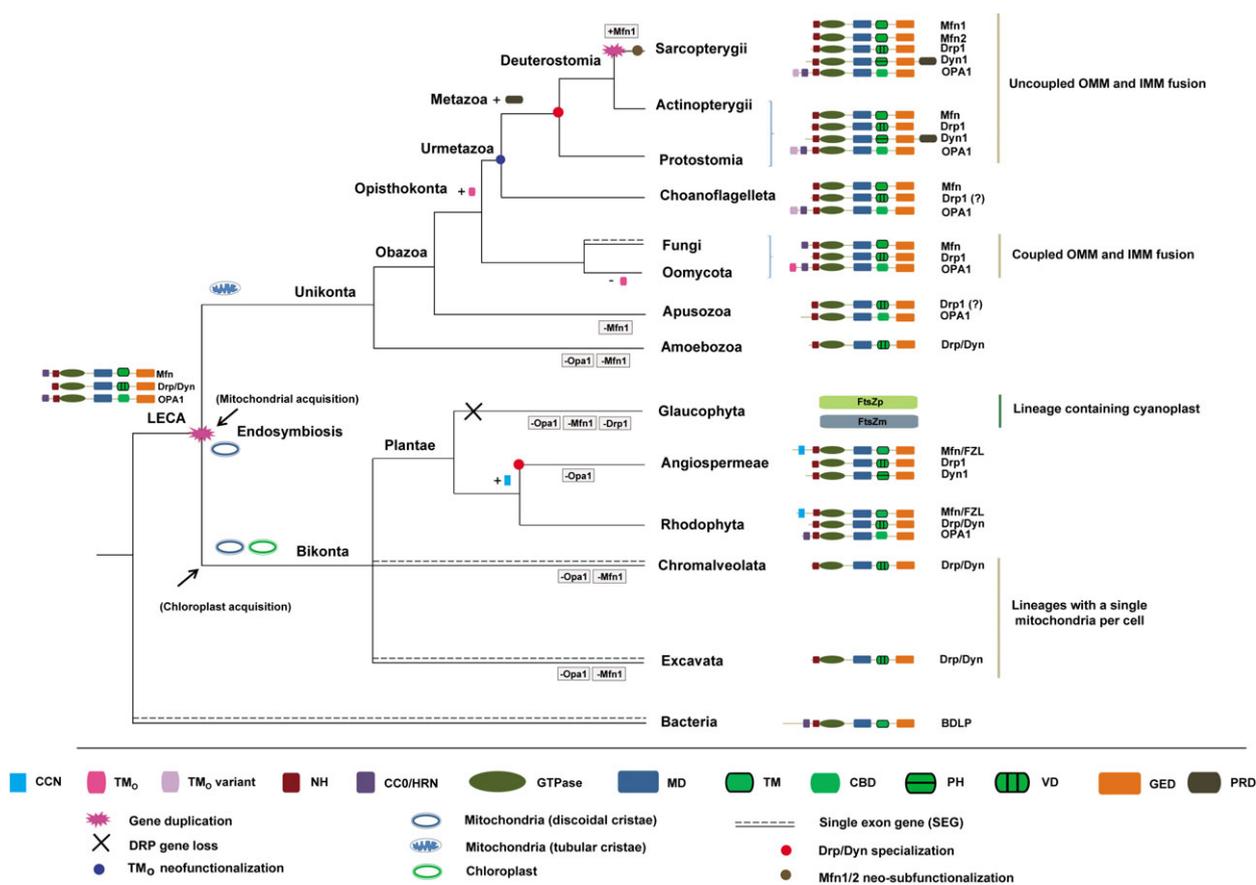


Fig. 5. Proposed evolutionary history of the mitochondrial fission and fusion dynamin-like proteins. The LECA with primordial mitochondria possessed a full complement of fission and fusion DSPs that were retained or lost subsequently in a lineage-specific manner. The endosymbiotic events leading to mitochondrial and chloroplast acquisition by eukaryotes is marked. The identity of the DSP and its domain architecture for each lineage is shown in the right column. Gene losses are indicated by – symbol along with the gene name. The domain addition events specific to each node is shown by + symbol. The X symbol represents loss of all canonical DRP paralogs in Glaucophytes. FtsZm and FtsZp represent the functionally equivalent proteins in Glaucophytes. The branch shown as a single line correspond to lineages containing intron coding genes, while the double dotted line represents lineages with single exon genes (SEG). Mfn/FZL orthologs localize to the chloroplast. The bifunctional fission DSP (mediating mitochondrial and vesicle scission) is represented by Drp/Dyn, while the specialized mitochondrial and vesicular DSPs are represented as Drp1 and Dyn1, respectively. Uncharacterized/putative fission DSP is shown as Drp1(?).

plants. Since *G. sulphuraria* contains multiple mitochondria per cell and displays tubulated mitochondrial networks, the absence of a Mfn ortholog that can specifically mediate mitochondrial fusion is intriguing [79]. Interestingly, the organellar localization signal presequence (36 residues) at the N-terminal of the Mfn/FZL ortholog here was predicted as CTS by ChloroP (ChloroP score = 0.559, cTP = Y), whereas it was identified as MTS by TargetP (TargetP mTP score = 0.696, cTP score = 0.129). The ‘ambiguous’ peptide appears to be a dual-targeting signal that can direct the protein to both mitochondria and chloroplast. Presence of dual targeted proteins is well recognized in plants, including in early diverging lineages such as bryophytes (reviewed in [80]). It follows that the *G. sulphuraria* Mfn/FZL may indeed be targeted to both chloroplast and mitochondria to carry out appropriate fusion roles in either organelle. However, the location of this presequence is similar to that in OPA1 implying its targeting to the IMS and not to the OMM as required for mitofusin. One possibility is that this ortholog may be targeted specifically to the chloroplast, and that the identity of the true mitofusin in the rhodophyte is unknown, akin to that in seed plants. It is also known that prediction of dual localization is inaccurate due to the limited availability of verified datasets and therefore, the MTS prediction here may be a false positive. Alternatively, albeit in a less likely scenario, the rhodophyte Mfn/FZL ortholog may mediate OMM fusion using a completely different mechanism. Considering that the green plant CTS sequences are clearly predicted by both programs (TargetP mTP score = 0.235, cTP score = 0.871, ChloroP score = 0.577, cTP = Y for *A. thaliana* Mfn/FZL), we conclude that chloroplast specificity of Mfn/FZL emerged independently in seed plants.

The absence of Mfn/FZL in the multicellular rhodophyte *C. crispus* can be attributed to the presence of the rhodoplast that contains a single lamellae with unstacked thylakoid, unlike the stacked thylakoids present in seed plants and unicellular rhodophyte *G. sulphuraria* [81]. In contrast, OPA1 is present in the multicellular *C. crispus*, whereas it is absent in the unicellular *G. sulphuraria* and *C. merolae*. Thus, the gain–loss events of fusion OPA1/Mfn paralogs in the rhodophytes appear to be strongly correlated with the occurrence of single/multiple mitochondrion per cell and the presence of unstacked/stacked thylakoids, in the multicellular and unicellular rhodophyte, respectively [79,82,83].

This work carried out to explore the origin and evolution of the eukaryotic mitochondrial fission–fusion apparatus also indicates that the BDLP is remarkably

similar to the mitofusin. It is recognized that the signature DSP domain architecture among the BDLPs is modified by deletions, fusion, and insertion of prokaryote-specific domains resulting in varied non-canonical forms that appear to have evolved to carry out divergent functional roles in membrane remodeling. Moreover, multiple paralogs distributed at distinct genomic locations or present as a tandem-pair/fused pair is not uncommon in bacteria [72,84,85]. Nevertheless, the divergence of BDLPs appears to have occurred independent of the diversification of eukaryotic DSPs. Since the function and molecular evolution of BDLPs in general is poorly understood, an independent evolutionary analysis of BDLPs together with the eukaryotic homologs may provide insights into the origin of the dynamin superfamily.

In summary, the data suggest that gene duplication of an ancestral DSP leading to the paralogs that mediate mitochondrial fusion and fission emerged concomitantly in the LECA. Perhaps, this machinery to mediate and regulate mitochondrial shape and interorganellar interactions enabled adaptation to subsequent physiological transitions. The patterns of divergence in gene gain–loss events, domain level architecture, coding regions, and gene structures appear to be function-selective and parallels the evolution of mitochondrial shape, ultrastructure, and metabolism across eukaryotic lineages. An alternate less parsimonious scenario assumes that the fusion and fission proteins in the lineages considered here emerged out of multiple independent duplication events in different eukaryotic lineages. Under this scenario, the fission DSP (Drp1 or Drp/Dyn) appeared in the LECA. The fusion DSP Mfn appeared independently in the bikont Angiospermeae and in the Urmetazoa in the unikonts, while the fusion DSP OPA1 emerged independently in the Obazoa and multicellular rhodophytes. We expect that the results presented here will pave the way for further functional characterization of the mitochondrial fission–fusion apparatus in eukaryotes.

Acknowledgements

We gratefully acknowledge an anonymous reviewer for the critical review and important comments and suggestions. The authors acknowledge financial support from IIT Madras. Infrastructure support from the Bioinformatics Infrastructure Facility funded by the Department of Biotechnology, Government of India, and the High Performance Computing Environment, IIT Madras, is acknowledged. We acknowledge Gopalakrishna Aradhyaam for fruitful discussions and Infant Sagayaraj Ravhe for help in phylogenetic tree analyses.

Data accessibility

Research data pertaining to this article is located at figshare.com: <https://dx.doi.org/10.6084/m9.figshare.7819535>.

Author contributions

SS performed the experiments and wrote the manuscript. SS and NM conceived the idea. NM designed and supervised the experiments, analyzed data, and edited the manuscript.

References

- Sagan L (1967) On the origin of mitosing cells. *J Theoret Biol* **14**, 225–275.
- Mackenzie S and McIntosh L (1999) Higher plant mitochondria. *Plant Cell* **11**, 571–585.
- Gray MW (2012) Mitochondrial evolution. *Cold Spring Harb Perspect Biol* **4**, 1–16.
- Gray MW, Burger G and Lang F (1999) Mitochondrial evolution. *Science* **283**, 1476–1481.
- Gray MW (2014) The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. *Cold Spring Harb Perspect Biol* **6**, 1–13.
- Franz Lang B, Gray MW and Burger G (1999) Mitochondrial genome and the origin of eukaryotes. *Annu Rev Genet* **33**, 351–397.
- Gawryluk RMR, Kamikawa R, Stairs CW, Silberman JD, Brown MW and Roger AJ (2016) The earliest stages of mitochondrial adaptation to low oxygen revealed in a novel Rhizarian. *Curr Biol* **26**, 2729–2738.
- Noguchi F, Shimamura S, Nakayama T, Yazaki E, Yabuki A, Hashimoto T, Inagaki Y, Fujikura K and Takishita K (2015) Metabolic capacity of mitochondrion-related organelles in the free-living anaerobic Stramenopile *Cantina marsupialis*. *Protist* **166**, 534–550.
- Makiuchi T and Nozaki T (2014) Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. *Biochimie* **100**, 3–17.
- Tomkins JL and Brown GS (2004) Population density drives the local evolution of a threshold dimorphism. *Nature* **431**, 1099–1103.
- Buczek D, Wojtkowska M, Suzuki Y, Sonobe S, Nishigami Y, Antoniewicz M, Kmita H and Makalowski W (2016) Protein import complexes in the mitochondrial outer membrane of Amoebozoa representatives. *BMC Genom* **17**, 99.
- Praefcke GJ and McMahon HT (2004) The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat Rev Mol Cell Biol* **5**, 15.
- Daumke O and Praefcke GJ (2016) Invited review: mechanisms of GTP hydrolysis and conformational transitions in the dynamin superfamily. *Biopolymers* **105**, 580–593.
- Daumke O and Roux A (2017) Mitochondrial homeostasis: How do dimers of mitofusins mediate mitochondrial fusion? *Curr Biol* **27**, R353–R356.
- Zhang J, Liu X, Liang X, Lu Y, Zhu L, Fu R, Ji Y, Fan W, Chen J, Lin B, *et al.* (2017) A novel ADOA-associated OPA1 mutation alters the mitochondrial function, membrane potential, ROS production and apoptosis. *Sci Rep* **7**, 5704.
- Zhu X, Perry G, Smith MA and Wang X (2013) Abnormal mitochondrial dynamics in the pathogenesis of Alzheimer's disease. *J Alzheimers Dis* **33**, 253–262.
- Cervený KL, Tamura Y, Zhang Z, Jensen RE and Sesaki H (2007) Regulation of mitochondrial fusion and division. *Trends Cell Biol* **17**, 563–569.
- Chen H and Chan DC (2009) Mitochondrial dynamics—fusion, fission, movement, and mitophagy—in neurodegenerative diseases. *Hum Mol Genet* **18**, 169–176.
- Frohlich C, Grabiger S, Schwefel D, Faelber K, Rosenbaum E, Mears J, Rocks O and Daumke O (2013) Structural insights into oligomerization and mitochondrial remodelling of dynamin 1-like protein. *EMBO J* **32**, 1280–1292.
- Heymann JA and Hinshaw JE (2009) Dynamins at a glance. *J Cell Sci* **122**, 3427–3431.
- Reddy PH, Reddy TP, Manczak M, Calkins MJ, Shirendeb U and Mao P (2011) Dynamin-related protein 1 and mitochondrial fragmentation in neurodegenerative diseases. *Brain Res Rev* **67**, 103–118.
- Schwaller B (2010) Cytosolic Ca²⁺ buffers. *Cold Spring Harb Perspect Biol* **2**, a004051.
- Kamthan A, Kamthan M, Kumar A, Sharma P, Ansari S, Thakur SS, Chaudhuri A and Datta A (2015) A calmodulin like EF hand protein positively regulates oxalate decarboxylase expression by interacting with E-box elements of the promoter. *Sci Rep* **5**, 14578.
- Bach D, Pich S, Soriano FX, Vega N, Baumgartner B, Oriola J, Daugaard JR, Lloberas J, Camps M, Zierath JR *et al.* (2003) Mitofusin-2 determines mitochondrial network architecture and mitochondrial metabolism. A novel regulatory mechanism altered in obesity. *J Biol Chem* **278**, 17190–17197.
- Chen H, Detmer SA, Ewald AJ, Griffin EE, Fraser SE and Chan DC (2003) Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. *J Cell Biol* **160**, 189–200.
- Dhingra R and Kirshenbaum LA (2014) Regulation of mitochondrial dynamics and cell fate. *Circ J* **78**, 803–810.
- Huang P, Galloway CA and Yoon Y (2011) Control of mitochondrial morphology through differential interactions of mitochondrial fusion and fission proteins. *PLoS ONE* **6**, e20655.
- van der Bliek AM, Shen Q and Kawajiri S (2013) Mechanisms of mitochondrial fission and fusion. *Cold Spring Harb Perspect Biol* **5**, a011072.

- 29 Zorzano A and Claret M (2015) Implications of mitochondrial dynamics on neurodegeneration and on hypothalamic dysfunction. *Front Aging Neurosci* **7**, 1–16.
- 30 Belenguer P and Pellegrini L (2013) The dynamin GTPase OPA1: more than mitochondria? *Biochim Biophys Acta* **1833**, 176–183.
- 31 Song Z, Ghochani M, McCaffery JM, Frey TG and Chan DC (2009) Mitofusins and OPA1 mediate sequential steps in mitochondrial membrane fusion. *Mol Biol Cell* **20**, 3525–3532.
- 32 Cao YL, Meng S, Chen Y, Feng JX, Gu DD, Yu B, Li YJ, Yang JY, Liao S, Chan DC *et al.* (2017) MFN1 structures reveal nucleotide-triggered dimerization critical for mitochondrial fusion. *Nature* **542**, 372–376.
- 33 Qi Y, Yan L, Yu C, Guo X, Zhou X, Hu X, Huang X, Rao Z, Lou Z and Hu J (2016) Structures of human mitofusin 1 provide insight into mitochondrial tethering. *J Cell Biol* **215**, 621–629.
- 34 Yan L, Qi Y, Huang X, Yu C, Lan L, Guo X, Rao Z, Hu J and Lou Z (2018) Structural basis for GTP hydrolysis and conformational change of MFN1 in mediating membrane fusion. *Nat Struct Mol Biol* **25**, 233–243.
- 35 Purkanti R and Thattai M (2015) Ancient dynamin segments capture early stages of host-mitochondrial integration. *PNAS* **112**, 2800–2805.
- 36 Robert X and Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.
- 37 Tamura K, Stecher G, Peterson D, Filipinski A and Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- 38 Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W and Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307–321.
- 39 Letunic I and Bork P (2007) Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128.
- 40 Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O and Pupko T (2012) FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res* **40**, W580–W584.
- 41 Pupko T, Péer I, Shamir R and Graur D (2000) A fast algorithm for joint maximum likelihood reconstruction of ancestral amino acid sequences. *Mol Biol Evol* **17**, 890–896.
- 42 Hu B, Jin J, Guo AY, Zhang H, Luo J and Gao G (2015) GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **31**, 1296–1297.
- 43 Gu X (1999) Statistical methods for testing functional divergence after gene duplication. *Mol Biol Evol* **16**, 1664–1674.
- 44 Ramachandran R and Schmid SL (2018) The dynamin superfamily. *Curr Biol* **28**, 411–416.
- 45 Wienke DC, Knetsch ML, Neuhaus EM, Reedy MC and Manstein DJ (1999) Disruption of a dynamin homologue affects endocytosis, organelle morphology, and cytokinesis in *Dictyostelium discoideum*. *Mol Biol Cell* **10**, 225–243.
- 46 Miyagishima SY (2003) A plant-specific dynamin-related protein forms a ring at the chloroplast division site. *Plant Cell Online* **15**, 655–665.
- 47 Miyagishima SY, Kuwayama H, Urushihara H and Nakanishi H (2008) Evolutionary linkage between eukaryotic cytokinesis and chloroplast division by dynamin proteins. *Proc Natl Acad Sci USA* **105**, 15202–15207.
- 48 Nishida K, Takahara M, Miyagishima SY, Kuroiwa H, Matsuzaki M and Kuroiwa T (2003) Dynamic recruitment of dynamin for final mitochondrial severance in a primitive red alga. *Proc Natl Acad Sci USA* **100**, 2146–2151.
- 49 Morgan GW, Goulding D and Field MC (2004) The single dynamin-like protein of *Trypanosoma Brucei* regulates mitochondrial division and is not required for endocytosis. *J Biol Chem* **279**, 10692–10701.
- 50 Gaechter V, Schraner E, Wild P and Hehl AB (2008) The single dynamin family protein in the primitive protozoan *Giardia Lamblia* is essential for stage conversion and endocytic transport. *Traffic* **9**, 57–71.
- 51 Wu YCR, Rasmussen MD and Kellis MD (2012) Evolution at the subgene level: domain rearrangements in the *Drosophila* phylogeny. *Mol Biol Evol* **29**, 689–705.
- 52 Marsh JA and Teichmann SA (2010) How do proteins gain new domains? *Gen Biol* **11**, 7–126.
- 53 Lynch M and Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155.
- 54 Bjorklund AKE, Light D, Frey-Skott S and Elofsson JA (2005) Domain rearrangements in protein evolution. *J Mol Biol* **353**, 911–923.
- 55 Okamoto PM, Herskovits JS and Vallee RB (1997) Role of the basic, proline-rich region of dynamin in Src homology 3 domain binding and endocytosis. *J Biol Chem* **272**, 11629–11635.
- 56 Krishnan S, Collett M and Robinson PJ (2015) SH3 domains differentially stimulate distinct dynamin I assembly modes and G domain activity. *PLoS ONE* **10**, e0144609.
- 57 Chappie JS and Dyda F (2013) Building A fission machine-structural insights into dynamin assembly and activation. *J Cell Sci* **126**, 2773–2784.
- 58 De Vecchis D, Cavellini L, Baaden M, Henin J, Cohen MM and Taly A (2017) A membrane-inserted structural model of the yeast mitofusin Fzo1. *Sci Rep* **7**, 10217.

- 59 Huang X, Zhou X, Hu X, Joshi AS, Guo X, Zhu Y, Chen Q, Prinz WA and Hu J (2017) Sequences flanking the transmembrane segments facilitate mitochondrial localization and membrane fusion by mitofusins. *Proc Natl Acad Sci USA* **114**, 9863–9872.
- 60 Gao H, Sage TL and Osteryoung KW (2006) FZL, an FZO-like protein in plants, is a determinant of thylakoid and chloroplast morphology. *Proc Natl Acad Sci USA* **103**, 6759–6764.
- 61 Anand R, Wai T, Baker MJ, Kladt N, Schauss AC, Rugarli E and Langer T (2014) The i-AAA protease YME1L and OMA1 cleave OPA1 to balance mitochondrial fusion and fission. *J Cell Biol* **204**, 919–929.
- 62 Cipolat S, Martins de Brito O, Dal Zilio B and Scorrano L (2004) OPA1 requires mitofusins 1 to promote mitochondrial fusion. *Proc Natl Acad Sci USA* **101**, 15927–15932.
- 63 DeVay RM, Dominguez-Ramirez L, Lackner LL, Hoppins S, Stahlberg H and Nunnari J (2009) Coassembly of Mgm1 isoforms requires cardiolipin and mediates mitochondrial inner membrane fusion. *J Cell Biol* **186**, 793–803.
- 64 Fülöp SG, Enyedi B, Vařlořnai P and Spät A (2011) The effect of OPA1 on mitochondrial Ca²⁺ signaling. *PLoS ONE* **6**, 11.
- 65 Pan R, Jones AD and Hu J (2014) Cardiolipin-mediated mitochondrial dynamics and stress response in *Arabidopsis*. *Plant Cell* **26**, 391–409.
- 66 Ehses S, Raschke I, Mancuso G, Bernacchia A, Geimer S, Tondera D, Martinou JC, Westermann B, Rugarli EI and Langer T (2009) Regulation of OPA1 processing and mitochondrial fusion by m-AAA protease isoenzymes and OMA1. *J Cell Biol* **187**, 1023–1036.
- 67 Naotada Ishihara YF, Oka T and Mihara K (2006) Regulation of mitochondrial morphology through proteolytic cleavage of OPA1. *EMBO J* **25**, 2966–2977.
- 68 Olichon A, Emorine LJ, Descoins E, Pelloquin L, Brichese L, Gas N, Guillou E, Delettre C, Valette A, Hamel CP *et al.* (2002) The human dynamin-related protein OPA1 is anchored to the mitochondrial inner membrane facing the inter-membrane space. *FEBS Lett* **523**, 171–176.
- 69 Wong ED, Wagner JA, Gorsich SW, Michael McCaffery J, Shaw JM and Nunnari J (2000) The dynamin-related GTPase, Mgm1p, is an intermembrane space protein required for maintenance of fusion competent mitochondria. *J Cell Biol* **151**, 341–352.
- 70 Herlan M, Bornhøvd C, Hell K, Neupert W and Reichert AS (2004) Alternative topogenesis of Mgm1 and mitochondrial morphology depend on ATP and a functional import motor. *J Cell Biol* **165**, 167–173.
- 71 Schwefel D, Frohlich C, Eichhorst J, Wiesner B, Behlke J, Aravind L and Daumke O (2010) Structural basis of oligomerization in septin-like GTPase of immunity-associated protein 2 (GIMAP2). *Proc Natl Acad Sci USA* **107**, 20299–20304.
- 72 Liu J, Noel JK and Low HH (2018) Structural basis for membrane tethering by a bacterial dynamin-like pair. *Nat Commun* **9**, 3345.
- 73 Babenko VN, Rogozin IB, Mekhedov SL and Koonin EV (2004) Prevalence of intron gain over intron loss in the evolution of paralogous gene families. *Nucleic Acids Res* **32**, 3724–3733.
- 74 Rogozin IB, Sverdlov AV, Babenko VN and Koonin EV (2005) Analysis of evolution of exon–intron structure of eukaryotic genes. *Brief Bioinform* **6**, 118–134.
- 75 Logsdon JM Jr, Stoltzfus A and Doolittle WF (1998) Molecular evolution: recent cases of spliceosomal intron gain? *Curr Biol* **8**, 560–563.
- 76 Lynch M (2002) Intron evolution as a population-genetic process. *Proc Natl Acad Sci USA* **99**, 6118–6123.
- 77 Coulombe-Huntington J and Majewski J (2007) Characterization of intron loss events in mammals. *Genome Res* **17**, 23–32.
- 78 Kiefel BR, Gilson PR and Beech PL (2004) Diverse eukaryotes have retained mitochondrial homologues of the bacterial division protein FtsZ. *Protist* **155**, 105–115.
- 79 Kuroiwa T, Nishida K, Yoshida Y, Fujiwara T, Mori T, Kuroiwa H and Misumi O (2006) Structure, function and evolution of the mitochondrial division apparatus. *Biochim Biophys Acta* **1763**, 510–521.
- 80 Murcha MW, Kmiec B, Kubiszewski-Jakubiak S, Teixeira PF, Glaser E and Whelan J (2014) Protein import into plant mitochondria: signals, machinery, processing, and regulation. *J Exp Bot* **65**, 6301–6335.
- 81 Margulis L and Chapman MJ (eds) (1999) Chapter two – KINGDOM PROTOCTISTA. In *Kingdoms and Domains* (Fourth Edition), pp. 117–230. Academic Press, San Diego, CA.
- 82 Lee RE (2008) *Phycology*. Cambridge University Press, Cambridge, UK.
- 83 Suzuki K, Kawano S and Kuroiwa T (1994) Single mitochondrion in acidic hot-spring alga: behaviour of mitochondria in *Cyanidium caldarium* and *Galdieria sulphuraria* (Rhodophyta, Cyanidiophyceae). *Phycologia* **33**, 298–300.
- 84 Bohuszewicz O, Liu J and Low HH (2016) Membrane remodelling in bacteria. *J Struct Biol* **196**, 3–14.
- 85 Burmann F, Ebert N, van Baarle S and Bramkamp M (2011) A bacterial dynamin-like protein mediating nucleotide-independent membrane fusion. *Mol Microbiol* **79**, 1294–1304.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. List of DSP sequences used in the study.

Table S2. Subcellular localization of DSP sequences.

Table S3. Sites displaying type II functional divergence across OPA1 and Mgm1 orthologs in the TM_O domain.

Table S4. Sites displaying type I and type II functional divergence across Mfn1 and Mfn2 orthologs identified from tetrapods.

Fig. S1. Summary of methodology.

Fig. S2. Ribbon representation of DSP structures.

Fig. S3. Phylogeny of DSPs.

Fig. S4. Functional divergence of tetrapod Mfn1 and Mfn2 in-paralogs.

Fig. S5. Weblogos representing the domain-wise consensus of all sequences from all phylogenetic groups.

Fig. S6. Schematic representing domain-wise presence of orthologous splice sites across protein sequences.

Fig. S7. Gene structure schematics of DSP sequences.

Fig. S8. Structural comparisons of MD and GED helical segments constituting the neck region of DSPs.

Text S1. Evolution and diversification of mitofusins.

File S1. Multiple sequence alignment (371 sequences) with splice site positions marked in yellow highlight.

File S2. Bayesian tree for the GTPase domain region (Newick format).

File S3. ML tree for the GTPase domain using structure-based sequence alignment (Newick format).

File S4. List of sites displaying type I and type II divergence between groups a) euteleostome Mfn2 vs. Tetrapod Mfn1, b) Metazoan Opa1 vs. Fungal Mgm1.

File S5. ML tree reconstructed from multiple sequence alignment of full-length proteins (380 nonredundant sequences) (Newick format).

File S6. Multiple sequence alignment of ATL and GBP sequences with Dyn1, Drp1, Mfn1, and Opa1 orthologs. Splice site positions are marked in yellow highlight.