# Myelin Tetraspan Family Proteins but No Non-Tetraspan Family Proteins Are Present in the Ascidian (*Ciona intestinalis*) Genome

ROBERT M. GOULD[1,2,*], HILARY G. MORRISON[2,3], EDWIN GILLAND[2], AND
ROBERT K. CAMPBELL[3,4]

[1] *Department of Anatomy and Cell Biology, University of Illinois at Chicago, Chicago, Illinois 60612;*
[2] *Marine Biological Laboratory, Woods Hole, Massachusetts 02543;* [3] *Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, Massachusetts 02543; and* [4] *Serono Reproductive Biology Institute, One Technology Place, Rockland, Massachusetts 02370*

**Abstract.** Several of the proteins used to form and maintain myelin sheaths in the central nervous system (CNS) and the peripheral nervous system (PNS) are shared among different vertebrate classes. These proteins include one-to-several alternatively spliced myelin basic protein (MBP) isoforms in all sheaths, proteolipid protein (PLP) and DM20 (except in amphibians) in tetrapod CNS sheaths, and one or two protein zero ($P_0$) isoforms in fish CNS and in all vertebrate PNS sheaths. Several other proteins, including 2′,3′-cyclic nucleotide 3′-phosphodiesterase (CNP), myelin and lymphocyte protein (MAL), plasmolipin, and peripheral myelin protein 22 (PMP22; prominent in PNS myelin), are localized to myelin and myelin-associated membranes, though class distributions are less well studied. Databases with known and identified sequences of these proteins from cartilaginous and teleost fishes, amphibians, reptiles, birds, and mammals were prepared and used to search for potential homologs in the basal vertebrate, *Ciona intestinalis.* Homologs of lipophilin proteins, MAL/plasmolipin, and PMP22 were identified in the *Ciona* genome. In contrast, no MBP, $P_0$, or CNP homologs were found. These studies provide a framework for understanding how myelin proteins were recruited during evolution and how structural adaptations enabled them to play key roles in myelination.

## Introduction

All modern-day gnathostomes, from mammals to elasmobranches, have structurally and compositionally similar myelin sheaths. Because these sheaths are absent in agnathans (Bullock *et al.,* 1984) and invertebrate chordates, myelination most likely developed in the common gnathostome ancestor. Myelin sheaths surround large-caliber axons of the central nervous system (CNS) and the peripheral nervous system (PNS) in a segmental fashion, forming internodes. These internodes, bordered axon membranes with high concentrations of sodium channels, promote saltatory conduction—a process that dramatically increases the effectiveness of neuronal communication. The evolution of saltatory conduction could have supported improved predation, escape, and other behaviors that offered advantages for the development and survival of the early gnathostome ancestor. The evolution of myelination required ancestral glial cells to target to growing axons, separate single axon segments from neighbors, and form insulating multilayered membrane sheaths of appropriate dimensions for each large-caliber axon. The final step, myelin formation, involved an up-regulation of specific glial cell proteins that function to cause the expanding plasma membrane to spirally wrap into a tightly compacted multilayered myelin sheath.

The aim of this study was to identify invertebrate homologs of vertebrate glial proteins adapted to shape myelin sheath formation. Focus was on the common sea squirt *Ciona intestinalis* (L.), an invertebrate chordate derived from a common ancestor of extant vertebrates. This animal was chosen because its entire genome is known (Dehal *et al.,* 2002) and the ascidian tadpole larva is emerging as an important model of the development of the chordate nervous system (Meinertzhagen *et al.,* 2004; Satoh, 2004; Passamaneck and Di Gregorio, 2005). Furthermore, the ascidian genome retains simplicity lost in the rounds of gene duplication associated with vertebrate evolution.

This study focused on structurally diverse low-molecular-weight proteins present in most vertebrate myelin sheaths. Included are compositionally prominent (20%–50% of protein mass) myelin basic protein (MBP), protein zero ($P_0$), and proteolipid protein (PLP) and "myelin-related" 2′, 3′-cyclic nucleotide 3′-phosphodiesterase (CNP), peripheral myelin protein 22 (PMP22), myelin and lymphocyte protein (MAL), and plasmolipin. Deletion, overexpression, or mutation of CNP, PMP22, and MAL all result in defects in myelination and/or myelin sheath stability, which further demonstrates their importance in myelination (*e.g.,* Lappe-Siefke *et al.,* 2003; Schaeren-Wiemers *et al.,* 2004).

## Materials and Methods

Initial myelin protein sequences were identified by text searches of the Entrez Protein database ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein⟩. Additional family members were identified using these sequences as queries for BLASTP searches of GenBank ⟨http://www.ncbi.nlm.nih.gov/BLAST/⟩. From the databases obtained ⟨http://www.mbl.edu/BiologicalBulletin/VIDEO/BB.video.html⟩, query sequences were selected (Table 1) and used in TBLASTN searches of release 1.0 of the *Ciona intestinalis* genome database ⟨http://genome.jgi-psf.org/cgi-bin/runAlignment?db=ciona4⟩. Gene model sequences matching the queries with expectation value less than 0.1 were then used as queries in reciprocal BLASTP searches of the GenBank non-redundant database. Annotations in the *Ciona* model, the presence of conserved protein domains (Marchler-Bauer and Bryant, 2004), and relatedness to subject myelin proteins in the reciprocal BLASTP searches were used as the basis for accepting or rejecting model sequences as being myelin protein family homologs. When queries failed to identify suitable gene models, searches were expanded to the *Ciona intestinalis* EST database (⟨http://ghost.zool.kyoto-u.ac.jp/indexrl.html⟩, TBLASTN). Sequences obtained from these searches were then used as queries for both GenBank and *Ciona* genome database searches. Complementing this approach and to achieve additional sensitivity, "jump

start" with PSI-BLAST (Altschul *et al.,* 1997) of the NCBI non-redundant (nr) protein database ⟨ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz⟩ was used to create a position-specific score matrix for each myelin protein family. "Jump-start" PSI-BLAST with vertebrate protein alignments containing the query sequences was performed as described in the documentation for BLAST. The alignments used for the "jump start" were built with ClustalX (Chenna *et al.,* 2003). PSI-BLAST was then run against the nr database for five iterations (including from each round the sequence matched with an E-value less than 0.001). The resulting family-specific scoring matrices were then used with the same PSI-BLAST queries to search release 1.0 of the *Ciona intestinalis* protein predictions ⟨http://www.jgi.doe.gov/⟩ for additional family members. The resulting sequences originating from EST and PSI-BLAST searches were considered true positives if supported by reciprocal BLAST matches to vertebrate proteins from the target families, and by matches to the Conserved Domain Database.

For each gene model identified, prediction of transmembrane helices ⟨http://www.cbs.dtu.dk/services/TMHMM-2.0/⟩ and presence of potential signal peptide cleavage sites ⟨http://www.cbs.dtu.dk/services/SignalP/⟩ were determined.

## Results

*Of prominent myelin proteins, only PLP family members are present in* Ciona

*Search for MBP homologs.* Myelin basic proteins (MBPs) are abundant in the myelin sheaths of gnathostome central nervous system (CNS) and peripheral nervous system (PNS) (Waehneldt *et al.,* 1986) and are a logical starting point for this study. Of the 55 MBP protein sequences from 16 species (9 mammals, chick, frog, zebrafish, pufferfish, and 4 elasmobranchs) that were identified, 8 were chosen as query sequences (Table 1). Because of complexity introduced by extensive alternative splicing and alternative promoter use (Campagnoni and Campagnoni, 2004), queries of the *Ciona* genome database were conducted with the largest MBP sequences available from each chosen species. No hits were obtained using the expectation cutoff of 0.1. Although relaxing this criteria (to E-value <1) resulted in four candidate matches, reciprocal BLAST and sequence analyses showed that none of these candidates were related to MBP. Similarly, candidate matches identified in the *Ciona intestinalis* EST database (TBLASTN, 5′ and 3′-ESTs) failed to yield reciprocal BLAST matches to MBP and lacked recognizable MBP domain sequences. Likewise, PSI-BLAST searches failed to identify MBP homologs. These efforts suggest that MBP homologs do not exist in *Ciona intestinalis.*

## Table 1

*Proteins used as queries to search the* Ciona *databases*

| Protein | Species | Accession # | # of amino acids | Similarity (%) |
|---|---|---|---|---|
| **Major myelin proteins** | | | | |
| MBP | Human | HBHUB | 197 | |
| | Bovine | MBBOB | 169 | 80.3 |
| | Mouse | P04370 | 195 | 92.5 |
| | Chick | NP_990611 | 174 | 70.1 |
| | Frog | BAA9174 | 176 | 63.1 |
| | Zebrafish | Ejf33a07 | 140 | 22.6 |
| | Pufferfish | CAG12843 | 217 | 27.6 |
| | Horned shark | CAA35761 | 155 | 46.3 |
| $P_0$ | Human | JHO252 | 248 | |
| | Rat | NP_058723 | 248 | 96.0 |
| | Chick | P37301 | 249 | 83.1 |
| | Frog | AAH41735 | 229 | 38.6 |
| | Trout | AAB34399 | 202 | 56.5 |
| | Horned shark | P20938 | 246 | 64.0 |
| PLP | Human | P60201 | 277 | |
| DM-20 | Mouse | AAA95516 | 242 | 86.3 |
| PLP | Chick | CAA43839 | 277 | 96.4 |
| PLP | Frog | S31491 | 280 | 84.3 |
| PLP | Lungfish | BAB16472 | 266 | 63.2 |
| Dm_1 | Zebrafish | AAL59885 | 245 | 64.3 |
| DM_ | Spiny dogfish | P36963 | 245 | 68.2 |
| DM | Fruit fly | AA69516 | 249 | 39.3 |
| **Myelin-related proteins** | | | | |
| CNP | Human | JC1517 | 421 | |
| | Bovine | NP_851336 | 400 | 93.5 |
| | Chick | NP_990381 | 420 | 80.0 |
| | Bullfrog | BAA24224 | 403 | 71.0 |
| | Pufferfish | FRUP00000141237 | 373 | 52.7 |
| gRICH | Goldfish | AAA84448 | 411 | 39.5 |
| Mal | Human | NP_002362 | 153 | |
| | Mouse | NP_034892 | 153 | 94.8 |
| | Chick | X415004 | 155 | 70.7 |
| | Frog | AAH77032 | 159 | 67.3 |
| | Zebrafish | NP_001004625 | 159 | 63.8 |
| Plasmolipin | Human | NP_057077 | 182 | |
| | Mouse | NP_080661 | 182 | 95.1 |
| | Chick | XP_414016 | 174 | 74.7 |
| | Frog | AAH73074 | 170 | 54.3 |
| | Pufferfish | CAF98911 | 145 | |
| PMP22 | Human | JN0503 | 160 | |
| | Mouse | P16646 | 161 | 93.2 |
| | Chick | XP_415596 | 160 | 87.0 |
| | Frog | AAH41727 | 160 | 86.9 |
| | Pufferfish | CAG06584 | 158 | 77.5 |
| | Zebrafish | AAH62526 | 157 | 69.4 |

The species, accession number, number of amino acids, and percent similarity to the human sequence are shown for each of the sequences used in the searches. More information can be found in the supplemental information (http://www.mbl.edu/Biologica/Bulletin/VIDEO/BB.video.html).

*Search for protein zero homologs.* Protein zero ($P_0$) is highly abundant (>50% of total protein) in fish CNS myelin sheaths and in almost all vertebrate PNS myelin sheaths (Kirschner *et al.,* 2004). Surprisingly, $P_0$ was not identified in zebrafish PNS (Brosamle and Halpern, 2002; Schweitzer *et al.,* 2003). Of 52 $P_0$ sequences from 12 species (6 mammals, chick, frog, 3 teleost fishes, and horned shark), 6 were chosen as queries (Table 1). Three produced hits to seven different scaffolds in the *Ciona* genome database, and five of these (identified with human and rat queries) matched the same 30-amino-acid region (161–190 in human $P_0$) that includes the single transmembrane (TM) domain (residues 153–179). Although these sea squirt sequences were 33%–40% identical and 70%–74% similar to rat and human $P_0$ sequences in this region, reciprocal GenBank searches with the model sequences did not suggest that these *Ciona* sequences were $P_0$ homologs. Likewise, reciprocal searches of GenBank with the two *Ciona* gene model sequences found with shark $P_0$ as query indicated that these were not $P_0$ homologs. Extending the analyses to the *Ciona intestinalis* EST database and with PSI-BLAST both failed to identify $P_0$ homologs in sea squirt.

*Search for lipophilin homologs.* Proteolipid protein (PLP) did not become the dominant CNS myelin protein until it appeared in amphibian CNS (Schliess and Stoffel, 1991; Yoshida and Colman, 1996; Venkatesh *et al.,* 2001). Yet, a more ancestral origin of this family is supported by identification of smaller DM20 and related M6 proteins in elasmobranches, teleosts, and lobe-finned fishes. Together, PLP, DM20, and M6 tetraspan proteins are grouped into a single family of lipophilin proteins (Gow, 1997). Lipophilin genes have already been identified in the insects—fruit fly, *Drosophila melangaster,* and silkworm, *Bombyx mori*— (Stecca *et al.,* 2000), indicating that the basic family predates the evolutionary divergence of deuterostomes and protostomes. Sequences from seven species (Table 1) were used to search the *Ciona* genome database and identified three possible matches, including two previously predicted gene models (ci0100152946—sea squirt DM-1 and ci0100149653—sea squirt DM-2; Table 2). An additional candidate (ci0100150877—DM-3) was obtained from the PSI-BLAST search (Table 2). The protein (sea squirt DM-3), derived from eight exons, was related to M6B protein in its N-terminal half (residues 9–195 of 441 amino acids). Sea squirt DM-1 is 36.3% similar to sea squirt DM-2 in overlapping regions (alignment contains four gaps in sea squirt DM-1) and 35.4% similar to sea squirt DM-3 (also with four gaps in sea squirt DM-1). Sea squirt DM-2 is only 21.1% similar to sea squirt DM-3 (entire sequences were aligned). None of these models contained the PLP-specific exon 3b.

To further characterize the relatedness of the three sea

**Table 2**

*Summary of sea squirt gene model sequences related to myelin proteins*

| Sea squirt protein | A. Properties of the sea squirt gene model | | | | | B. Results of BLASTP GenBank searches with sea squirt models | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Gene model | Size | Annotated hit | Top hit EST | Cluster | Conserved domain | Region | Top GenBank hit | Accession # | Expect |
| DM-1 | ci0100152946 | 259 | Myelin PLP | cign043j24 | CLSTR37034r1 | Myelin_PLP | 1–243 | Mouse M6B | P35803 | 5e-47 |
| DM-2 | ci0100149653 | 433 | Myelin PLP | cinc030e24 | CLSTR00983r1 | Myelin_PLP | 1–229 | Chick M6b | XP_416830 | 2e-19 |
| DM-3 | ci0100150877 | 441 | Hypothetical | ciht034b02 | CLSTR10664r1 | Myelin_PLP | 9–196 | Mouse M6B | AAF87987 | 9e-10 |
| PL-1 | ci0100134996 | 162 | None/plasmolipin[1] | cibd077c03 | CLSTR14342r1 | MARVEL | 23–151 | Mouse CLF-8[2] | XP_510988 | 9e-17 |
| PL-2 | ci0100141316 | 120 | Hypothetical | ciem829k02 | CLSTR44564r1 | MARVEL | 2–120 | Chimp CLF-4[3] | XP_523385[4] | 1e-14 |
| PMP22-1 | ci0100153863 | 165 | PMP22/EMP/MP20 | ciem846m10 | CLSTR01413r1 | PMP22_claudin | 12–155 | Fugu Hypothetical | CAG08900 | 6e-11 |
| PMP22-2 | ci0100153716 | 227 | PMP22/EMP/MP20 | cinc003m04 | CLSTR02642r1 | PMP22_claudin | 36–163 | Frog MGC68719 | AAH60425 | 1e-05 |

Relationships between sea squirt gene models identified in this study and vertebrate proteins. A. Gene models identified from searches of the *Ciona intestinalis* database, including model identification number, number of amino acids, annotation in the *Ciona* search result panel, the top EST identified, and the gene cluster with which the model is associated. B. BLASTP search results, including conserved domains, amino acid range that contained the conserved domain, the top search hit, and the expectancy value. *Abbreviation:* CLF—chemokine-like factor super family protein.

[1] Although the gene model was not annotated in the JGI search result, it was annotated as plasmolipin in the corresponding EST database search.
[2] Chimp and rat plasmolipin proteins were the 5th and 6th top hit (9e-16 and 6e-15).
[3] Zebrafish PMP-22 proteins were identified as 9th top hit (6e-08).
[4] Zebrafish PMP-22 proteins were identified as 3rd top hit (1e-04).

squirt sequences to known lipophilin proteins, a multiple-sequence alignment was made which included all human, mouse, chick, frog, lungfish, zebrafish, and spiny dogfish lipophilin proteins (Appendix Fig. A1). This alignment indicated that sea squirt DM-3 was far less related to known vertebrate lipophilin proteins than either sea squirt DM-1 or DM-2. A simpler alignment that included sea squirt DM-1 and DM-2 and several vertebrate proteins indicated that DM-1 was more "vertebrate-like" (Fig. 1). This conclusion was further supported by the observation that sea squirt DM-1 has intron-exon boundaries identical to those in PLP/DM20 (Fig. 1), while the boundaries in sea squirt DM-2 are totally different (not shown). The four TM domains predicted for both sea squirt DM-1 and DM-2 (DM-3 had only three predicted TM domains, all in the N-terminal half) fell close to the positions identified in other lipophilin proteins (not shown). Also, the relative sizes of loops linking the putative TM regions were similar for sea squirt DM-1 and DM-2 models and vertebrate lipophilin sequences. When analyzed for signal peptide cleavage, sea squirt DM-1, like human DM-20, had a predicted signal peptide at its N-termini; sea squirt DM-2 and DM-3 did not.

Alignments of human, chick, frog, lungfish, zebrafish, and spiny dogfish lipophilin proteins revealed that all these proteins shared 41 amino acids (Appendix Fig. A1). Sea squirt DM-1 had 27 of these amino acids and sea squirt DM-2 had 21. A large number were cysteine residues. Of four cysteine residues in human DM20 that are fatty acylated, three are present in sea squirt DM-1 sequence and two

**Figure 1.** Alignment of model sequences for sea squirt DM-1 and DM-2s (includes the region from amino acids 44–380 of the DM-2 model sequences) with sequences for human DM20 (AAA60118), human M6A1 (NP 005268), human M6B1 (AAC28560), frog DMβ (AAK17942), and spiny dogfish Dmβ (AAC59640). Alignments were made with the Clustal W program using the Vector NTI Version 9.0 (Invitrogen, Inc.). The amino acid positions are in parentheses in front of each sequence, and those of the overall alignments are shown at the top. The alignment is coded as follows: (1) White letters on a black background represent identity across the alignment. (2) White letters on a gray background represent blocks of identity. (3) Black letters on a gray background represent blocks with conserved replacements. (4) Black letters on a white background represent weak or no similarity. Four transmembrane (TM) domains as depicted in a PLP/DM/M6 sequence alignment (fig. 1, Hudson, 2004) are indicated just below the alignment ([—]) and closely match locations of predicted TM domains in both sea squirt DM-1 and DM-2. The six intron/exon borders common to sea squirt DM-1 and human PLP (ENSG00000123560) are marked with upward arrows (↑) below the aligned sequences. Sites of the four cysteine residues used in fatty acylation are indicated above the aligned sequences (white letters on a black background), as are the four cysteine residues that participate in disulfide bond formation (white letters on a gray background). Other common cysteine residues are indicated on the line above the sequences. Those present in M6 proteins and not in DM20/PLP are marked with gray shading. The lysine residue involved in PLP glycosylation is marked above the sea squirt DM-1 sequence (black letter with gray background).

```
                                 1                C  CC  C                   C              C C 60
Sea squirt DM-1      (1)   ------------------MGCFDSCLRCLS---LVPWASLIAAVLCWAGTALFCGTAHE
Sea squirt DM-2s     (1)   HPATTVATTPKRKKVRYRDAGCCDGCVNCLITIGTMPEATAFAVILCWLGTAMFIGGWS
Human DM20           (1)   ------------------MGLLECCARCIV---GAPEASLVATGLCFFGVALFCGCGHE
Human M6A1           (1)   --------MEENMEEGQTQKGCFECCIKCLG---GIPYASLIATILLYAGVALFCGCGHE
Human M6B1           (1)   MKPAMETAAEENTEQSQERKGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCGHV
Frog DMβ             (1)   --------MEENMEEGQTQKGCFECCIKCLG---GIPYASLIATILLYAGVALFCGCGHE
Spiny dogfish DMβ    (1)   --------MEENMEEGQTQKGCFECCIKCLG---GIPYASLIATILLYAGVALFCGCGHE
TM/Intron-Exon                                ↑              [-----------I-----------]


                                 1                                                      120
Sea squirt DM-1      (39)  AVSNSYILLNSTDFYWN--------------------------ETSTD-TIIQNFKYS
Sea squirt DM-2s     (61)  AIDGTMNLWAQFGPMAGNPYNHPPPQYLYTGNALTPQFINLNRDDLGNRVPAVFQAIKYS
Human DM20           (39)  ALTGTEKLIETYFSKN-YQ-----------------------DYEYLINVIHAFQYV
Human M6A1           (50)  ALSGTVNILQTYFEMARTAG----------------------DTLDVFTMIDIFKYV
Human M6B1           (58)  ALAGTVAILEQHFSTN-AS-----------------------DHALLSEVIQLMQYV
Frog DMβ             (50)  ALTGTVNILQTYFEMARTAG----------------------DTLDVFTMIDIFKYV
Spiny dogfish DMβ    (50)  ALSGTVTILQNNFEVVRGAG----------------------DTIDVFTMIDIFKYV
TM/Intron-Exon                                                              ↑   [--II-


                                 121              C        C                       180
Sea squirt DM-1      (70)  IYGIAAFMFVLSILLFVDGILSTRVVKSDYEEGCKTSCCGLCIGIFYTMLVYICGVGWLL
Sea squirt DM-2s     (121) FYGLIPFMFVFTIMLACDNRHSSRAMASKKRS-CKSSSSGICLTAALIGVTYILILAWVF
Human DM20           (72)  IYGTASFFFLYGALLLAEGFYTTGAVRQIFGD-YKTTICGKGLSATFVGITYALTVVWLL
Human M6A1           (85)  IYGIAAAFFVYGIILMVEGFFTTGAIKDLYGD-FKITTCGRCVSAWFIMLTYLFMLAWLG
Human M6B1           (91)  IYGIASFFFLYGIILLAEGFYTTSAVKELHGE-FKTTACGRCISGMFVFLTYVLGVAWLG
Frog DMβ             (85)  IYGIAAAFFVYGIILLMVEGFFTTGAIKDLYGD-FKITTCGRCVSAWFIMLTYLFMLAWLG
Spiny dogfish DMβ    (85)  IYGVAAAFFVYGIILLMVEGFFTTGAIKDLYGD-FKITTCGRCVSGWFIMLTYIFMLAWLG
TM/Intron-Exon               ----II--------]                        ↑     [-------III-


                                 181            C              C                  240
Sea squirt DM-1      (130) ITCFAAFPIHFFIVANSQCKAN--------------TCIDFRQTGILPAHL----PDDIG
Sea squirt DM-2s     (180) FLAFTTLGVYYYRISMLRCSDLQ-EREFSTGVL-KDVCIDIVQLGVMMFKNTQDNGFGKL
Human DM20           (131) VFACSAVPVYIYFNTWTTCQSIAFP-SKTSAS-IGSLCADARMGVLPWN------AFPG
Human M6A1           (144) VTAFTSLPVYMYFNLWTICRNTTLVEG-------ANLCIDLRQFGIVTIGE-----EKKI
Human M6B1           (150) VFGFSAVPVFMFYNIWSTCEVIKSPQTNGTTG-VEQICVDIRQYGIIPWN------AFPG
Frog DMβ             (144) VTAFTSLPVYMYFNLWTICRNATLIEE-------ANFCIDLRQFGIVTVGE-----EKKL
Spiny dogfish DMβ    (144) VTAFTSLPVFMYFNIWTLCQNVTIMES-------TDLCFDLRQFGIVPIHE-----QKTV
TM/Intron-Exon               ----III------]                              ↑


                                 C242       K  C                              300
Sea squirt DM-1      (172) RICGDDLTKFCSTP----TFNVTYQLYIVAFAGAAITLLSMKQFLMCLSANFAYLKMTKK
Sea squirt DM-2s     (238) CGPGDSMQSYGKLEGYCQNYYVVYQYMLVAFSGCLLNIIAMINFIMILTINYNDLARKRA
Human DM20           (183) KVCGSNLLSICKTA----EFQMTFHLFIAAFVGAAATLVSLITFMIAATYNFAVLKLMGR
Human M6A1           (192) CTVSENFLRMCEST----ELNMTFHLFIVALAGAGAAVIAMVHYLMVLSANWAYVKDACR
Human M6B1           (203) KICGSALENICNTN----EFYMSYHLFIVACAGAGATVIALIIYMMATTYNYAVLKFKSR
Frog DMβ             (192) CSPSDNFLRMCEST----ELNMTFHLFIVALAGAGAAVIAMVHYLMVLSANWAYVKDACR
Spiny dogfish DMβ    (192) CTLNENFSKLCQSN----DLNMTFHLFIVALAGAGAAVIAMVHYLMVLSANWAYVKDACR
TM/Intron-Exon               ↑           [-----------IV--↑-----------]


                                 1                             332
Sea squirt DM-1      (228) IAVYEDTKYREEMELNDIINTARSNERLTYKY
Sea squirt DM-2s     (298) APTRSTASAAPSQASTTPIV-ARSVRSIPNTS
Human DM20           (239) GTKF----------------------------
Human M6A1           (248) MQKYEDIKSKEEQELHDIHST-RSKERLNAYT
Human M6B1           (259) EDCCTKF-------------------------
Frog DMβ             (248) MQKYEDIKSKEEQELHDIHST-RSKERLNAYT
Spiny dogfish DMβ    (248) MQKYEDIKSKEEQELHDIHST-RSKERLNAYT
```

are present in sea squirt DM-2 (Fig. 1). All four cysteine residues used for disulfide bonding in human DM20 are conserved in sea squirt DM-1, whereas one residue is absent in sea squirt DM-2 (Fig. 1). Three cysteine residues, shared among M6 proteins, sea squirt DM-1, and DM-2, are absent in PLP and DM20 sequences and may reflect differences between these families. Consistent with this observation, neighbor-joining phylogenetic trees (not shown) indicate that sea squirt DM-1 and DM-2 are more similar to M6A proteins than they are to M6B or DM20. Alignments of these evolutionarily diverse lipophilin proteins suggest that cysteine residues are important in providing stability in the proteins' membrane environments (Hudson, 2004). The conservation of other amino acids, both within and outside the TM domains, suggests a focus for future studies to determine their functional importance.

Although the 35-amino-acid PLP-specific exon 3b sequence is limited to tetrapods, a search of the *Ciona* genome database (expectation cutoff, E <10) identified six scaffolds, one of which (174, no gene model) had a 17-amino-acid sequence (KGRRSRGQHKGHDLRHL) 59% identical to exon 3b, and another, adjacent to the first (746, no gene model), with seven of nine (RVCHCKGRW) amino acids identical to exon 3b. BLASTP searches for short, nearly exact matches of GenBank with both these sequences identified PLP exon 3b sequences as best hits. Surprisingly, a similar search of the pufferfish genome database ⟨http://genome.jgi-psf.org/cgi-bin/runAlignment?db=fugu6⟩ did not find a sequence nearly as similar to the PLP-specific exon 3b as the sea squirt sequences. Further genome organization studies might provide insight into whether these *Ciona* sequences were co-opted in the transformation of a DM/M6 protein into PLP.

The expression patterns of sea squirt DM-1 and sea squirt DM-2 were inferred from EST database matches to the gene models (Table 2). The results suggest that DM-1 is expressed in gastrulae and neurulae and in the neural complex and that sea squirt DM-2 is more broadly distributed throughout larval development. Tissue expression for DM-2 includes muscle, notochord, epidermis, endoderm, endodermal strand, nerve cord, brain, papilla, neuron, mesenchyme, trunk lateral cells, and trunk ventral cells. *In situ* hybridization studies needed to determine whether expression includes neuronal and ependymal cells were not reported.

*Among myelin-related proteins, MAL/plasmolipin and PMP22, but not CNP, have* Ciona *homologs*

*Search for CNP/RICH homologs.* 2′, 3′-Cyclic nucleotide 3′-phosphodiesterase (CNP) is a protein long associated with CNS and PNS myelination (see Sprinkle, 1989). Although it is believed to be absent in fish myelin, we identified a gene model (FRUP00000141237) in the pufferfish

database that encodes a 373-amino-acid protein with 52% similarity to human and bullfrog CNP. In addition, a more distantly related protein, regeneration-induced goldfish CNP homolog (gRICH), was identified in regenerating optic nerve (Wilmot *et al.,* 1993). gRICH has structural and functional properties similar to those of CNP, mainly in the C-terminal two-thirds of the molecule (Ballestero *et al.,* 1997). A zebrafish homolog of gRICH has also been reported (Ballestero *et al.,* 1999). Database searches identified 45 sequences related to CNP/gRICH, including genes from four mammals, chick, two amphibians, pufferfish, and goldfish. Searches of the JGI *Ciona* genome with five CNP and one gRICH sequences (Table 1) identified scaffolds only with the gRICH sequence. Reciprocal BLASTP searches of GenBank with the gRICH gene model sequence hits failed to identify CNP, gRICH proteins, or conserved domains related to these proteins. A gene model (ci0100130899, 1138 amino acids) related to CNP, though not to gRICH sequences, was identified from searches of the *Ciona* EST database and through a PSI-BLAST search. The only relationship of this protein to CNP was the presence of an ATPase A-box domain (Braun *et al.,* 2004) that occurs in CNP, but not gRICH, proteins.

*MAL/plasmolipin homologs in the* Ciona *genome.* The myelin and lymphocyte protein (MAL) is a tetraspan protein distinct from the PLP/DM20 family expressed by oligodendrocytes and Schwann cells (Bronstein, 2000; Frank, 2000). This protein likely participates in targeting of lipids and proteins that associate with lipid rafts to myelin and with stabilizing compact myelin (Schaeren-Wiemers *et al.,* 2004). MAL similarly targets raft-related lipids and proteins in kidney cells (Martin-Belmonte *et al.,* 2000). Of 37 MAL sequences identified in database searches, 5 were used as queries (Table 1). The chick and frog sequences were previously annotated in the databases as MAL proteins; zebrafish was not. A possible sea squirt MAL protein was identified with frog MAL as query (scaffold 20, ci0100134996). A reciprocal BLASTP search with this model sequence identified a conserved MARVEL domain (characteristic of MAL) and matches to MAL family members including plasmolipin (Magyar *et al.,* 1997) and chemokine-like factor super family (CLF)-8 proteins (Han *et al.,* 2003). Models from other *Ciona* hits were found to be unrelated to MAL family proteins in the reciprocal BLASTP searches.

Like MAL, plasmolipin proteins are expressed by both oligodendrocytes and Schwann cells (Gillen *et al.,* 1996) and are enriched in CNS and PNS myelin sheaths (Fischer *et al.,* 1994). Plasmolipin proteins (human, mouse, chick, pufferfish, and frog) are 34%–41% similar to MAL proteins from the same species. Whereas the chick sequence was previously annotated as a plasmolipin, the frog and puffer-

fish sequences are listed as hypothetical proteins. These sequences were less similar to human plasmolipin than the corresponding frog and pufferfish MAL and PMP22 proteins were to human counterparts (Table 1). Queries with human, mouse, and frog plasmolipin sequences identified the gene model identified in the MAL protein search (above) and several other gene models that were unrelated to MAL/plasmolipin proteins. The single hit with the pufferfish sequence as query identified the same protein (ci0100141316) that was identified in our PSI-BLAST search with MAL sequences. This model contained a MARVEL domain over its entire length and identified CLF family-4 proteins as top hits in reciprocal BLASTP searches (Table 2). A third *Ciona* candidate identified only by PSI-BLAST (ci0100144565) lacked the MARVEL domain but yielded CLF-7 and CLF-4 proteins as top hits when used as a query in BLASTP against GenBank. Although the same size (163 amino acids) as other MAL and plasmolipin proteins, it has only weak similarity to them.

Alignments of the three sea squirt models with known MAL and plasmolipin sequences clearly demonstrated that sea squirt PL-1 is more "vertebrate-like" than the other *Ciona* proteins (not shown). Four predicted TM domains of sea squirt PL-1 and sea squirt PL-2 align closely with those identified in plasmolipin and MAL sequences (Fig. 2). Analyses for predicted signal peptide cleavage sites indicate differences among human MAL and plasmolipin and the two sea squirt models. Whereas both MAL and plasmolipin are predicted to have signal anchors, sea squirt PL-1 is predicted to be a non-secretory protein and sea squirt PL-2 is predicted to contain a signal peptide. At the level of gene structure, only two exons (separated in the middle of TMII) are predicted for sea squirt PL-1, whereas human plasmolipin contains four exons.

Alignment of all vertebrate plasmolipin and MAL proteins identifies 14 common amino acids (not shown). Sea squirt PL-1 contains 10 of these conserved residues, and sea squirt PL-2 contains 4. The phylogenetic tree (constructed with neighbor joining, Vector NTI, version 9.0, Invitrogen) indicates that sea squirt PL-1 clusters with plasmolipin and not with MAL proteins (not shown). Coupled with reciprocal BLASTP searches, the fact that sea squirt PL-1 has 18 amino acids in common with plasmolipin sequences (25 if the pufferfish sequence is not included) and only 13 amino acids in common with MAL sequences strengthens the assignment of the sea squirt sequence to the plasmolipin clade. Compared with the high content of cysteine residues (12–16) in DM/M6 proteins, plasmolipin and MAL proteins contain only 2–3 cysteine residues. Hypothetical zebrafish MAL contains four cysteine residues, and sea squirt PL-1 contains five cysteine residues, mostly in the predicted TM domains. Of eight amino acids at the start of TMII that are highly conserved among MAL family proteins (Frank,

2000), six are conserved in sea squirt PL-1 (Fig. 2). Seven of eight additional amino acids for MAL/plasmolipin/BENE proteins (Magyar *et al.,* 1997) are conserved in sea squirt PL-1 (Fig. 2).

The sea squirt PL-1 identified ESTs indicating plasmolipin as its best match (Table 2). Expression was highest in young adults and in blood cells, though larvae and neural complex hits were also recorded. ESTs related to sea squirt PL-2 identified a cluster with no information on expression or on localization.

*PMP22 homologs in the* Ciona *genome.* Peripheral myelin protein 22 (PMP22) is a tetraspan protein abundant in mammalian PNS myelin (Suter, 2004). It is currently widely studied, in part because mutations (duplications, deletions, and missense substitutions) are a leading cause of human neuropathies and because it has an important role in PNS myelination (Jetten and Suter, 2000). We conducted database searches with 6 of 35 identified sequences, including mammal, chick, frog, pufferfish, and zebrafish (Table 1). Human, mouse, pufferfish, and zebrafish queries recognized one candidate in *Ciona* (ci0100153863—PMP22-1), and the pufferfish and zebrafish sequence identified an additional candidate (ci0100153716—sea squirt PMP22-2). Both these sea squirt sequences are predicted to contain PMP22_claudin domains and recognized vertebrate PMP22 proteins among top hits in reciprocal BLAST analyses (Table 2). The sequences were 37.3% and 28.2% similar to human PMP22 and 28.8% similar to one another. An alignment was made of these sequences and known PMP22 sequences (Fig. 3). The predicted TM domains for the sea squirt PMP22-1 and PMP22-2 models align closely with the four TM domains predicted for human and frog PMP-22. All have a relatively large first extracellular loop. These loops, suggested to have importance for transmembrane interactions (Sanders *et al.,* 2001), are 37.2% (sea squirt PMP22-1) and 32.6% (sea squirt PMP22-2) similar to the extracellular loop in human PMP22. The C-terminal tail of sea squirt PMP22-2 is much larger (59 amino acids) than the three residue tails of sea squirt PMP22-1 and human and frog PMP22. Like the human, both sea squirt PMP22 models have three similarly sized exons located within the coding region (Fig. 3). Only 10 amino acids are common among the two sea squirt PMP22-related sequences and vertebrate PMP22 sequences. Whereas all PMP22 have four cysteine residues that are completely aligned in human, mouse, chick, and frog (three of these are aligned in the zebrafish and pufferfish sequences), only the first cysteine residues in the sea squirt models align with the vertebrate PMP22 sequences. Of 30 amino acids common among PMP22, EMP1, EMP2, and EMP-3 (Taylor and Suter, 1996), 8 are conserved in both sea squirt PMP22 sequences, and an additional one, at different sites, is conserved in each

```
                                1                                        S  LG          VL50
Sea squirt PL-1   (1)  -----------MESQTTTTTTTTTHTTTTVSLNVNYVTSPLGILRICEMVLG
Human PL          (1)  --MAEFPSKVSTRTSSPAQGAEASVSALRPDLGEVRSRLGAIMLLQLVLG
Mouse PL          (1)  --MAEFPSKVSTRTSSPAQGVGASVSALRPDLGFVRSALGVIALLQLALG
Chick PL          (1)  --MAGLPGAAGTRSSAP--GAAAILPSPRLDEAFLCSPLGGIMGAQAVLG
Frog PL           (1)  MEGERPRSDTVTTTVSSHMETISLGGSIAYDRSFLRSPTGVILLMEIMFG
Human MAL         (1)  ----------------MAPAAATGGSTLPSGFSVFTTLPDLLFIFEFIFG
Frog MAL          (1)  -----------MAATTPNAPVYEDTSSLPSGFRVLATFPDVLMIPDLVFG
TM/Intron-Exon                                              [-----I--↑--

                       L52 A                ********* W                      100
Sea squirt PL-1  (40)  LICWAIVASLSWLNYSPPWQFVMFVAVTSWITTVILFIIYVGAFHASICS
Human PL         (49)  LLVWALIADTPYHLYP-AYGWVMFVAVFLWLVTIVLFNLYIFQLHMKL-Y
Mouse PL         (49)  LLVWALIADTPYHLYP-AYGWVMFVAVFLWLVTIVFFIIYIFQLHMKL-Y
Chick PL         (47)  LLVWALLADTTYHLHA-AYGWVLFVSIFLWVVTVLLFVMYILQLPMKF-Y
Frog PL          (51)  LLVWALIAGSEYFLFS-AFGWVMFVAVFYWVLSVFFFLLHLTRANTRI-T
Human MAL        (35)  GLVWILVASSLVPWPL-VQGWVMFVSVFCFVATTTLIILYIIGAHG---G
Frog MAL         (40)  SLVWILVASSRVPFPI-VQGWVMYVSVSLFFFSVLLMLLYCARAHG---G
TM/Intron-Exon         -------]        [---------II---------]

                       101P*        F      V   *    *       *        *    *150
Sea squirt PL-1  (90)  GLPWAIIELLFNGFWAVFYFIAACVAA-AYTAG--------WSLLVAATV
Human PL         (97)  MVPWPLVLMIFNISATVLYITAFIACS-AAVDLTSLRGTRPYNQRAAASF
Mouse PL         (97)  MVPWPLVLLIFFVAATVLYITAFIACA-AAVDLTSLRGSRPYNQRSAASF
Frog PL          (99)  -KVPWSLVGLCFNGSAFVLYLIAAVVEA-SSVNKDVHQHHNYNSWTASSF
Chick PL         (95)  MVPWPLMLMIFNAVATVLYIAAFVTCS-AAVRPTSWR-QWDYNRRAAASF
Human MAL        (81)  ETSWVTLDAAYHCTAALFYLSASVLEALATITMQDGFTYRHYHENIAAVV
Frog MAL         (86)  KSSWTAVDAFYHCVAAYFYLSASVLESYSTIALSNVTIYQLYQENIAAVV
TM/Intron-Exon         [--↑------III--------]                          [--↑

                       *152       *                           187
Sea squirt PL-1 (131)  FAWFTTICYVVHAAFAFKEYRG-GSFSGSSVNT----
Human PL        (146)  FACLVMIAYGVSAFFSYQAWRGVGSNAATSQMAGGYA
Mouse PL        (146)  FACLVMIAYGVSAFFSFQAWRGVGSNAATSQMAGGYS
Chick PL        (143)  FAGLVMLTYGASTFFSFRA------------------
Frog PL         (147)  FAFIVTVCYALSTYFSFQAWRTKS-------------
Human MAL       (131)  FSYIATLLYVVHAVFSLIRWKSS--------------
Frog MAL        (136)  FAYLATLSYTIHAIFSLIRWKGSS-------------
TM/Intron-Exon         ------IV---------]
```

**Figure 2.** The top sea squirt model sequence for PL-1 was aligned with representative plasmolipin and myelin and lymphocyte protein (MAL) sequences, using the Clustal W program. Codes for alignment identities and conservative replacements are described in Figure 1. Sequence numbering is the same as in Figure 1. The four predicted transmembrane (TM) domains in human plasmolipin sequence, close to those reported for rat plasmolipin (Gillen *et al.,* 1996), are indicated as in Figure 1. Three intron-exon boundaries marked in the human plasmolipin sequence (ENST00000219207, http://www.ensembl.org/) are indicated. Amino acids conserved among plasmolipin/MAL/BENE proteins (Frank, 2000) are marked with asterisks (*) above the sea squirt PL-1 sequence, and amino acids shared between squirt PL-1 and known plasmolipin, but not MAL, sequences are indicated by white letters on a gray background above the sea squirt sequence.

sequence. Like the zebrafish PMP-22 sequence, neither of the sea squirt PMP 22 sequences retains the N-glycosylation site present in mammalian PMP22 (Fig. 3). The hypothetical pufferfish PMP22 contains this site. Extension of these

studies with PSI-BLAST identified 10 additional models that were related to claudin family proteins, though not specifically to PMP22. A phylogenetic tree for members of the PMP22/EMP3/MP20 family (Fig. 4) shows that the sea

```
                  1   M   LL          H    #   L    T      ↓ W      G        {W NC  50
Sea squirt PMP22-1 (1)  MVCTKKIQIASAAIALIGFIFLLVAICTNYWITSDVESNSGLWRACSAG-
Sea squirt PMP22-2 (1)  -MGLGGCEAFGIVLSLVGTALLLTAQLLAPWKTVGD-TWIGIWRDCNVGR
Human PMP22        (1)  ---MLLLLLSIIVLHVAVLVLLFVSTIVSQWIVGNG-HATDLWQNCSTSS
Mouse PMP22        (1)  ---MLLLLLGILFLHIAVLVLLFVSTIVSQWLVGNG-HTTDLWQNCTTSA
Chick PMP22        (1)  ---MLLLLLGIIVLHVTVLVLLFVSTIVSQWLVNGG-QTADLWQNCTSG-
Frog PMP22         (1)  ---MLLLLLGIIVLHIAILVLLFVSTIVSSWLVGNG-YSADLWQNCSGP-
Pufferfish PMP22   (1)  ---MLLLLLGIILLHIAALVILFVSTIVSAWTWSET-STSDLWMNCTTI-
Zebrafish PMP22    (1)  ---MLVILFATIVLHLTSLIILFIATCASAWRTHGD-TSYDLWYDCTQI-
TM/Intron-exon          [---------I---------]                      ↑
```
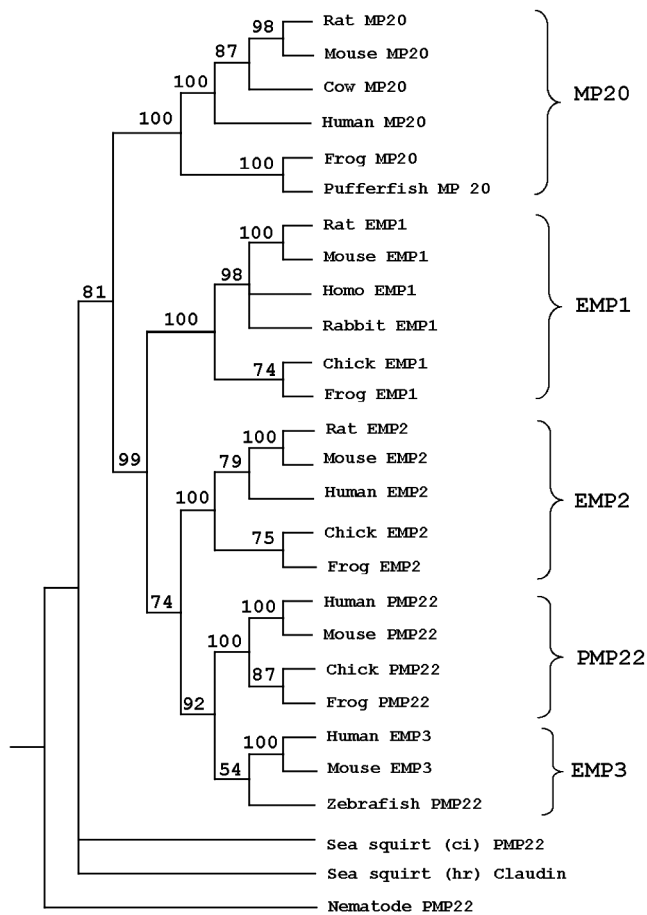
```
                  51       C          ↓ W     Q   M LS I    }      CQL     G  F
Sea squirt PMP22-1 (50) NCENLKIGSSRKYENYEMVRGFGVLAVCVCFLGIVLSLLSCVIRRVKIGP
Sea squirt PMP22-2 (49) LDQCGGLTDWLKRIDVTAARGCFIAADVLFLISILMLFAICADKKGGLI
Human PMP22        (47) SGNVHHCFSSSPNEWLQSVQATMILSIIFSILSLFLFFCQLFTLTKGGRF
Mouse PMP22        (47) LGAVQHCYSSSVSEWLQSVQATMILSVIFSVLALFLFFCQLFTLTKGGRF
Chick PMP22        (46) TGAIFQCLTSSTNEWLQSVQAMMILSVIFSVLSLFLFFCQLFTLTKGGRF
Frog PMP22         (46) ATGTWQCLTSSSNEWLQSVQAMMILSIIFSVLSLFLFFCQLFTLTKGGRF
Pufferfish PMP22   (46) VGAGQRCDPASTGEWIQAVQALMILSIIFSCLSLFLFFCQLFTLQKGGRF
Zebrafish PMP22    (46) -NGGYKCSGNGDADWLQAVQALMVLATIFCLISFIIFLCQLFTLVKGGRF
TM/Intron-exon                 ↑              [---------II---------]
```

```
                  101           ↓  C        IY        W                   G150
Sea squirt PMP22-1 (100) SVVATLYLMAAFFGLTSLALFTSIVNTKNSD----------VNRAWGYSF
Sea squirt PMP22-2 (99)  TAKGGLEITAGIIILAGCGAFTALNLVDLGAISDIGLFGLELRRDFGYAF
Human PMP22        (97)  YITGIFQILAGLCVMSAAAIYTVRHPEWHLN----------SDYSYGFAY
Mouse PMP22        (97)  YITGFFQILAGLCVMSAAAIYTVRHSEWHVN----------TDYSYGFAY
Chick PMP22        (96)  YLTGIFQILAGLCVMSGAAIFTVRHTDWHEAS---------ENTSYGFAY
Frog PMP22         (96)  YLTGLFQVLAAVCVMAGASIFTARHTEWQTA----------EVASYGFAY
Pufferfish PMP22   (96)  FLTGTFQILASLFVMSGAIIYTVMSPEWVPQ----------INTYGYSY
Zebrafish PMP22    (95)  FFTAVFQILASLFVMSGAIIYTVMSPQWKND----------SDTYGYAF
TM/Intron-exon          [------↑--III--------]                         [--
```

```
                  151L W   F          #   Y    LRK               200
Sea squirt PMP22-1 (140) ILGWFAFPLTLFPGVVMTYVELKMTN-----------------------
Sea squirt PMP22-2 (149) YLGWASGGVVFISGICLVVGGNQRSKKEKASNSLPVVMIPQYGPPKDVNV
Human PMP22        (137) ILAWVAFPLALLSGVIYVILRKRE-------------------------
Mouse PMP22        (137) ILAWVAFPLALLSGIIYVILRKREL------------------------
Chick PMP22        (137) ILAWLAFPLALASGIIYVILRKRE-------------------------
Frog PMP22         (136) ILAWVAFPLAAISGIIYIILRKRE-------------------------
Pufferfish PMP22   (135) ILAWVAFPLALISGLIYVILRKRE-------------------------
Zebrafish PMP22    (134) VLAWLAFPLTLISGLIYIILRKKE-------------------------
TM/Intron-exon          ------IV----------]
```

```
                  201                     229
Sea squirt PMP22-2 (199) NTVSRNQVTYYSTASMRCINRYSIRRWRQ
```

**Figure 3.** Two sea squirt model sequences for peripheral myelin protein 22 (PMP22) were aligned with representative PMP22 sequences, using the Clustal W program. Codes for alignment identities are described in Figure 1. The four predicted transmembrane (TM) domains in human PMP22 are smaller than reported (Naef and Suter, 1998). They align closely with predicted TM domains determined for the two sea squirt models. The five conserved tyrosine residues in vertebrate PMP22 sequences are marked (shaded white W) above sea squirt PMP-22-1. The three that are also present in sea squirt PMP22-1 and PMP22-2 are shown with a black background. The four cysteine residues common among tetrapod PMP22 (three are common with fish sequences) are indicated by shaded white Cs above the sea squirt PMP22-1 sequence. The region conserved among PMP22 and connexin proteins that might be involved in extracellular interactions (Sanders *et al.,* 2001) is indicated by brackets {} above the alignment. Sea squirt model cysteine residues close to the second conserved cysteine are shaded within the alignment. The single N-glycosylation site, marked above the sea squirt PMP22-1 sequence, is not conserved in the sea squirt or the zebrafish sequences. Residues conserved among PMP22, EMP1, EMP2 and EMP3 are represented with letters above the sea squirt PMP22-1 sequence. The point mutations L→P (position 16) in trembler and G→D (position 150) in trembler-J are designated with the symbol # above the sea squirt PMP22-1 sequence.

**Figure 4.** Phylogenetic relationship among PMP22/EMP/MP20 homologs. The tree was obtained with the minimum-evolution algorithm using PAUP (Swofford, 2003), based on the amino-acid alignment of 27 selected proteins. The initial tree was calculated using the neighbor-joining method and tested with 100 bootstrap replicates. The bootstrap values, which represent a measure of confidence, are shown above each node that was supported in a majority of the replicates. Accession numbers for the sequences represented are in the order P47801, NP 001414.1, NP 036975.1, P54850, XP 416203.1, AAH74477.1, NP 031955.1, NP 001415.1, AAH81865.1, XP 414935.1, AAH44968.1, NP 001416.1, O35912, NP 446223.1, P55344, P56563, NP 776527.1, AAH79741.1, CAG03357.1, NP 509258.1, JN0503, P16646, XP415595.1, AAH41727.1, AAH62526.1, ciona4:153863, and CAB96130.1. *Abbreviations:* ci, *Ciona intestinalis;* hr, *Halocynthia roretzi.*

squirt sequences are outliers and do not associate with a specific family member, suggesting that these sequences predate the division of the family into the different protein subfamilies observed in vertebrates. Sea squirt PMP22-1 is associated with a cluster (Table 2) whose EST hits are concentrated in young adults and larvae and in heart. Sea squirt PMP22-2 is associated with a different cluster (Table 2) whose EST hits locate in the young adult and neural complex.

## Discussion

How myelination, a prominent component of all vertebrate nervous systems, evolved is an important question in biology. Myelination requires a complex series of interactions between neurons and glial cells to foster a nervous system characterized by rapid, stable, and efficient transmission of nerve impulses from neuronal soma to terminal specializations. Perhaps, as in other systems such as jaws (Shigetani *et al.,* 2002), the vertebrate endocrine system (Campbell *et al.,* 2004), and antibody-mediated immunity (Klein and Nikolaidis, 2005), preexisting pieces were connected into the myelination process by gnathostome ancestors. Although focus here is limited to the origins of a few glial proteins recruited for myelin assembly, one should realize that very many proteins are required both for the coordinated interactions between differentiating glial cell and growing axons that precede myelin sheath formation and for the formation of paranodal and nodal specializations. These are events that require precise targeting of axonal and glial lipids and proteins to specialized membrane domains where glial cells and neuron processes interact.

Here, we show that four proteins involved in myelin sheath formation and stabilization have homologs in the ascidian genome; these four are proteolipid protein (PLP), myelin and lymphocyte protein (MAL), plasmolipin, and peripheral myelin protein 22 (PMP22). We also find that three other proteins, including the two most important myelin proteins, myelin basic protein (MBP) and protein zero ($P_0$), do not. MBP and $P_0$ are abundant in all sheaths except when $P_0$ is replaced by PLP in tetrapod CNS sheaths. Furthermore, mutant animals that lack either protein have the poorest ability of all myelin protein mutants to form sheaths. Clearly, MBP-deficient shiverer and mld mutants are unable to form CNS myelin sheaths (Privat *et al.,* 1979), and $P_0$-knockout mice are unable to form PNS sheaths (Giese *et al.,* 1992). Mutants lacking all other myelin proteins—including PLP, MAL, and 2′, 3′-cyclic nucleotide 3′-phosphodiesterase (CNP)—make relatively normal myelin before instability, axon degeneration, or both set in (Boison and Stoffel, 1994; Lappe-Siefke *et al.,* 2003; Schaeren-Wiemers *et al.,* 2004). These results suggest the importance of finding the evolutionary origins of MBP and $P_0$.

All proteins required for myelin sheath formation and maintenance were either adopted from other situations or found additional niches when they evolved for roles in myelination. MBP is expressed in spleen and lymph node macrophages (Liu *et al.,* 2001), $P_0$ in embryonic Schwann cells (Bhattacharyya *et al.,* 1991), DM20 in oligodendrocyte progenitors (Dong *et al.,* 1999), M6 proteins in neurons (Yan *et al.,* 1993), CNP in lymphoid

tissues (Bernier *et al.,* 1987), MAL in lymphocytes (Alonso and Weissman, 1987), plasmolipin in kidney (Fischer and Sapirstein, 1994), and PMP22 in fibroblasts (Schneider *et al.,* 1988) and epithelial cells (Notterpek *et al.,* 2001). Furthermore, in several cases, for example, MBP (Kitamura *et al.,* 1990) and PMP22 (Bosse *et al.,* 1994), separate promoters were adapted for high-level expression in myelinating cells.

The present study employs a database mining approach, similar to one used by others (*e.g.,* Sasakura *et al.,* 2003) to identify members of select protein families in the *Ciona* genome. Here, focus is on proteins used for myelin sheath formation and maintenance. Restricting the search to low-molecular-weight, single-domain proteins and starting from databases that contained, where possible, representative sequences from mammals, birds, reptiles (most poorly represented), amphibian, and teleost and cartilaginous fish, we successfully identified eight sea squirt protein relatives of myelin lipophilins, MAL/plasmolipin, and PMP22 proteins. The best model, in terms of relatedness or expectancy value, protein structure, and gene organization, was sea squirt DM-1. It not only was identified with all queries tested, it also had the same intron-exon structure and the highest expectancy (Table 2) of all myelin homologs. Other lipophilin homologs, as well as MAL/plasmolipin and PMP22 models, were all identified from a restricted subset of query sequence searches. Judged on the basis of intron-exon structure and relatedness to known myelin proteins, these are less vertebrate-like. As previously proposed, reciprocal BLASTP searches, which identified both conserved domains and *bona fide* myelin protein homologs as top hits in database searches, were used as criteria to determine whether sea squirt models were in fact homologs to proteins of interest.

From this search for myelin protein homologs, it was particularly interesting that all sea squirt homologs were tetraspan family proteins. No *Ciona* homologs were identified for the three non-tetraspan proteins used in our search. In addition, olig1, a non-tetraspan protein that is an important transcription factor for myelination (Zhou *et al.,* 2000), was not found in a search of the *Ciona* genome (Satou *et al.,* 2003). Connexins, a tetraspan protein family that includes at least two "myelin-related" members—connexin-32 (Dermietzel *et al.,* 1997) and connexin-29 (Altevogt *et al.,* 2002)—are clearly abundant in sea squirt (Sasakura *et al.,* 2003), though homologs of the myelin-specific connexins were not identified. In addition to studies focused on *Ciona,* studies with other non-vertebrate genomes identified two other myelin tetraspan family proteins, namely PMP22/claudin in *Caenorhabditus elegans* (Agostoni *et al.,* 1999) and single DM/M6 proteins in *Drosophila melangaster* and *Bombyx mori* (Stecca *et al.,* 2000). With availability of complete fruit fly and nematode genomes, the absence of MBP, $P_0$,

and CNP homologs suggests that these proteins likely evolved relatively late. As additional genome projects become available, especially those for amphioxus and lamprey, it will be interesting to see if they contain MBP and $P_0$ homologs.

Further support for differences in evolution of myelin tetraspan *versus* non-tetraspan proteins comes from consideration of orthologs. The protein families that relate to MBP, $P_0$, and CNP (excluding gRICH) are comparatively small, consisting usually of a single member. Even with teleost fishes that have probably undergone an additional round of genome duplication (Jaillon *et al.,* 2004), only single MBP, $P_0$, and CNP proteins were identified. Either the properties of mutations caused their removal from the genome or they were fashioned relatively late. In striking contrast, database searches identify many different proteins that fall into lipophilin, MAL/plasmolipin, and PMP22 families.

Taken together, these findings suggest that genes encoding different tetraspan protein families, including lipophilin, MAL/plasmolipin, PMP22/claudin, and connexins, are ancient and were important in the functioning and survival of the common urochordate/chordate ancestor. They also indicate that ancestors of other key myelin proteins, including MBP, $P_0$, CNP, and olig1, most likely developed rather late, possibly with the first vertebrates. The need for different types of tetraspan family proteins can be realized not only from their presence in a basal vertebrate, but also by the numbers of family members present in different species today. There are several DM/M6 proteins in human (Werner *et al.,* 2001), amphibians (Yoshida *et al.,* 1999), and teleost and cartilaginous fishes. Pufferfish has three PMP22-like proteins, as well as many representatives of the larger claudin family (Loh *et al.,* 2004), whose members are particularly important for the function of tight junctions (Tsukita and Furuse, 2002). Further, the family extends to nematodes (Asano *et al.,* 2003).

As evident from the literature and from alignments shown here, tetraspan proteins used in myelin differ in structure (Bronstein, 2000). For example, whereas MAL/plasmolipin and PMP22 proteins have few cysteine residues, lipophilin proteins have many. Although additional studies are needed to determine whether they are products of posttranslational modification, as is the case with PLP/DM20, conserved locations are suggestive. Alignments show both similarities and differences between sea squirt proteins and vertebrate homologs. In most cases the changes are small enough to allow conserved domain sequences to be identified (Table 2). Clearly the common myelin_PLP, MARVEL, and PMP22_claudin domains that distinguish these proteins contain important keys to their functions.

All three sea squirt DM proteins have recognizable myelin_PLP domains. DM-1 is closest to the vertebrate-

like lipophilin family and appears closest to M6A on the basis of similarity in the C-terminal tail (Fig. 1). In BLASTP searches, all three DM models identify M6B proteins as top hits (Table 2). Studies of expression patterns of the mRNAs and the upstream regulatory regions in developing ascidians (Kusakabe, 2005) may provide clues as to how these proteins evolved to function in both neurons and glial cells. The identification of a potential sequence related to PLP-specific exon 3b may also prove useful in untangling the relationships that allowed myelin PLP to evolve. Cysteine residues and cysteine residue modifications clearly play a role in lipophilin protein function (Hudson, 2004), and efforts to see if these modifications exist in M6 family and in sea squirt DM proteins will not only strengthen the conviction that they are functionally important, but also open avenues to understanding how the modifications influence function.

The functions of MAL, plasmolipin, and PMP22 in myelination are not entirely clear, as the processes whereby proteins and lipids assemble to form and stabilize the expanding plasma membranes of oligodendrocytes and Schwann cells are complex. Growing interest in lipid raft and raft signaling in myelin formation (*e.g.,* Schafer *et al.,* 2004) and evidence that these proteins are included in rafts in other cells (*e.g.,* Hasse *et al.,* 2002) open up studies that will further our understanding of the role of different tetraspan proteins in the elaboration of myelin.

In summary, we have found that the pre-duplicated genome of the sea squirt, *Ciona intestinalis,* contains homologs of a subset of proteins that are important for myelination. We hope that further studies on the location, expression requirements, and structural properties of these proteins in sea squirt and other simpler nervous systems will aid in our understanding of myelin evolution.

## Acknowledgments

## Literature Cited

Agostoni, E., S. Gobessi, C. Brancolini, and C. Schneider. 1999.    Identification and characterization of a new member of the *gas3/PMP22* gene family in *C. elegans. Gene* **234:** 267–274.

Alonso, M. A., and S. M. Weissman. 1987.    cDNA cloning and sequence of MAL, a hydrophobic protein associated with human T-cell differentiation. *Proc. Natl. Acad. Sci. USA* **84:** 1997–2001.

Altevogt, B. M., K. A. Kleopas, F. R. Postma, S. S. Scherer, and D. L. Paul. 2002.    Connexin29 is uniquely distributed within myelinating glial cells of the central and peripheral nervous systems. *J. Neurosci.* **22:** 6458–6470.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997.    Gapped BLAST and PSI BLAST: a new database of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.

Asano, A., K. Asano, H. Sasaki, M. Furuse, and S. Tsukita. 2003.    Claudins in *Caenorhabditis elegans:* their distribution and barrier function in the epithelium. *Curr. Biol.* **13:** 1042–1046.

Ballestero, R. P., G. R. Wilmot, B. W. Agranoff, and M. D. Uhler. 1997.    gRICH68 and gRICH70 are 2′,3′-cyclic-nucleotide 3′-phosphodiesterases induced during goldfish optic nerve regeneration. *J. Biol. Chem.* **272:** 11479–11486.

Ballestero, R. P., J. A. Dybowski, G. Levy, B. W. Agranoff, and M. D. Uhler. 1999.    Cloning and characterization of zRICH, a 2′,3′-cyclic-nucleotide 3′-phosphodiesterase induced during zebrafish optic nerve regeneration. *J. Neurochem.* **72:** 1362–1371.

Bernier, L., F. Alvarez, A. Mentaberry, J. G. Schembri, D. D. Sabatini, and D. R. Colman. 1987.    Molecular cloning of a 2′,3′-cyclic nucleotide 3′-phosphodiesterase: mRNAs with different 5′ ends encode the same set of proteins in nervous and lymphoid tissues. *J. Neurosci.* **7:** 2703–2710.

Bhattacharyya, A., E. Frank, N. Ratner, and R. Brackenbury. 1991.    $P_0$ is an early marker of the Schwann cell lineage in chickens. *Neuron* **7:** 831–844.

Boison, D., and W. Stoffel. 1994.    Disruption of the compacted myelin sheath of axons of the central nervous system in proteolipid protein-deficient mice. *Proc. Natl. Acad. Sci. USA* **91:** 11709–11713.

Bosse, F., G. Zoidl, S. Wilms, C. P. Gillen, H. G. Kuhn, and H. W. Müller. 1994.    Differential expression of two mRNA species indicates a dual function of peripheral myelin protein *PMP22* in cell growth and myelination. *J. Neurosci. Res.* **37:** 529–537.

Braun, P. E., M. Gravel, and J. Lee. 2004.    2′,3′-Cyclic nucleotide 3′-phosphodiesterase: structure, biology, and function. Pp. 499–522 in *Myelin Biology and Disorders,* R. A. Lazzarini, J. W. Griffin, H. Lassmann, K.-A. Nave, R. H. Miller, and B. D. Trapp, eds. Elsevier Academic Press, Amsterdam.

Bronstein, J. M. 2000.    Function of tetraspan proteins in the myelin sheath. *Curr. Opin. Neurobiol.* **10:** 552–557.

Brosamle, C., and M. E. Halpern. 2002.    Characterization of myelination in the developing zebrafish. *Glia* **39:** 47–57.

Bullock, T. H., J. K. Moore, and R. D. Fields. 1984.    Evolution of myelin sheaths: both lamprey and hagfish lack myelin. *Neurosci. Lett.* **48:** 145–148.

Campagnoni, A. T., and C. W. Campagnoni. 2004.    Myelin basic protein gene. Pp. 387–400 in *Myelin Biology and Disorders,* Vol. 1, R. A. Lazzarini, J. W. Griffin, H. Lassmann, K.-A. Nave, R. H. Miller, and B. D. Trapp, eds. Elsevier Academic Press, Amsterdam.

Campbell, R. K., N. Satoh, and B. M. Degnan. 2004.    Piecing together evolution of the vertebrate endocrine system. *Trends Genet.* **20:** 359–366.

Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson. 2003.    Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31:** 3497–3500.

Dehal, P., Y. Satou, R. K. Campbell, J. Chapman, B. Degnan, A. De Tomaso, B. Davidson, A. Di Gregorio, M. Gelpke, D. M. Goodstein *et al.* 2002.    The draft genome of *Ciona intestinalis:* insights into chordate and vertebrate origins. *Science* **298:** 2157–2167.

Dermietzel, R., M. Farooq, J. A. Kessler, H. Althaus, E. L. Hertzberg, and D. C. Spray. 1997.    Oligodendrocytes express gap junction proteins connexin32 and connexin45. *Glia* **20:** 101–114.

Dong, Z. P., A. Sinanan, D. Parkinson, E. Parmantier, R. Mirsky, and K. R. Jessen. 1999.    Schwann cell development in embryonic mouse nerves. *J. Neurosci. Res.* **56:** 334–348.

Fischer, I., and V. S. Sapirstein. 1994.    Molecular cloning of plasmolipin: characterization of a novel proteolipid restricted to brain and kidney. *J. Biol. Chem.* **269:** 24912–24919.

Fischer, I., R. Durrie, and V. S. Sapirstein. 1994. Plasmolipin: the other myelin proteolipid. A review of studies on its structure, expression, and function. *Neurochem. Res.* **19:** 959–966.

Frank, M. 2000. MAL, a proteolipid in glycosphingolipid enriched domains: functional implications in myelin and beyond. *Prog. Neurobiol.* **60:** 531–544.

Giese, K. P., R. Martini, G. Lemke, P. Soriano, and M. Schachner. 1992. Mouse $P_0$ gene disruption leads to hypomyelination, abnormal expression of recognition molecules, and degeneration of myelin and axons. *Cell* **71:** 565–576.

Gillen, C., M. Gleichmann, R. Greiner-Petter, G. Zoidl, S. Kupfer, F. Bosse, J. Auer, and H. W. Müller. 1996. Full-length cloning, expression and cellular localization of rat plasmolipin mRNA, a proteolipid of PNS and CNS. *Eur. J. Neurosci.* **8:** 405–414.

Gow, A. 1997. Redefining the lipophilin family of proteolipid proteins. *J. Neurosci. Res.* **50:** 659–664.

Han, W., P. Ding, M. Xu, L. Wang, M. Rui, S. Shi, Y. Liu, Y. Zheng, Y. Chen, T. Yang, and D. Ma. 2003. Identification of eight genes encoding chemokine-like factor superfamily members 1–8 (*CKLFSF1–8*) by in silico cloning and experimental validation. *Genomics* **81:** 609–617.

Hasse, B., F. Bosse, and H. W. Müller. 2002. Proteins of peripheral myelin are associated with glycosphingolipid/cholesterol-enriched membranes. *J. Neurosci. Res.* **69:** 227–232.

Hudson, L. D. 2004. Proteolipid protein gene. Pp. 401–420 in *Myelin Biology and Disorders,* Vol. 1, R. A. Lazzarini, J. W. Griffin, H. Lassmann, K.-A. Nave, R. H. Miller, and B. D. Trapp, eds. Elsevier Academic Press, Amsterdam.

Jaillon, O., J.-M. Aury, F. Brunet, J.-L. Petit, N. Stange-Thomann, E. Mauceli, L. Boneau, C. Fischer, C. Ozouf-Costaz, A. Bernot *et al.* 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431:** 946–957.

Jetten, A. M., and U. Suter. 2000. The peripheral myelin protein 22 and epithelial membrane protein family. *Prog. Nucleic Acid Res. Mol. Biol.* **64:** 97–129.

Kirschner, D. A., L. Wrabetz, and M. L. Feltri. 2004. The P0 gene. Pp. 523–545 in *Myelin Biology and Disorders,* R. A. Lazzarini, J. W. Griffin, H. Lassmann, K.-A. Nave, R. H. Miller, and B. D. Trapp, eds. Elsevier Academic Press, Amsterdam.

Kitamura, K., S. L. Newman, C. W. Campagnoni, J. M. Verdi, T. Mohandas, V. W. Handley, and A. T. Campagnoni. 1990. Expression of a novel transcript of the myelin basic protein gene. *J. Neurochem.* **54:** 2032–2041.

Klein, J., and N. Nikolaidis. 2005. The descent of the antibody-based immune system by gradual evolution. *Proc. Natl. Acad. Sci USA* **102:** 169–174.

Kusakabe, T. 2005. Decoding cis-regulatory systems in ascidians. *Zool. Sci.* **22:** 129–146.

Lappe-Siefke, C., S. Goebbels, M. Gravel, E. Nicksch, J. Lee, P. E. Braun, I. R. Griffiths and K. A. Nave. 2003. Disruption of Cnp1 uncouples oligodendroglial functions in axonal support and myelination. *Nat. Genet.* **33:** 366–374.

Liu, H. B., A. J. Mackenzie-Graham, K. Palaszynski, S. Liva, and R. R. Voskuhl. 2001. "Classic" myelin basic proteins are expressed in lymphoid tissue macrophages. *J. Neuroimmunol.* **116:** 83–93.

Loh, Y. H., A. Christoffels, S. Brenner, W. Hunziker, and B. Venkatesh. 2004. Extensive expansion of the claudin gene family in the teleost fish, *Fugu rubripes. Genome Res.* **14:** 1248–1257.

Magyar, J. P., C. Ebensperger, N. Schaeren-Wiemers, and U. Suter. 1997. Myelin and lymphocyte protein (MAL/MVP17/VIP17) and plasmolipin are members of an extended gene family. *Gene* **189:** 269–275.

Marchler-Bauer, A., and S. H. Bryant. 2004. CD-Search: protein domain annotation on the fly. *Nucleic Acids Res.* **32:** W327–W331.

Martin-Belmonte, F., R. Puertollano, J. Millan, and M. A. Alonso. 2000. The MAL proteolipid is necessary for the overall apical delivery of membrane proteins in the polarized epithelial Madin-Darby canine kidney and Fischer rat thyroid cell lines. *Mol. Biol. Cell* **11:** 2033–2045.

Meinertzhagen, I. A., P. Lemaire, and Y. Okamura. 2004. The neurobiology of the Ascidian tadpole larva: recent developments in an ancient chordate. *Annu. Rev. Neurosci.* **27:** 453–485.

Naef, R., and U. Suter. 1998. Many facets of the peripheral myelin protein PMP22 in myelination and disease. *Microsc. Res. Tech.* **41:** 359–371.

Notterpek, L., K. J. Roux, S. A. Amici, A. Yazdanpour, C. Rahner, and B. S. Fletcher. 2001. Peripheral myelin protein 22 is a constituent of intercellular junctions in epithelia. *Proc. Natl Acad. Sci USA* **98:** 14404–14409.

Passameneck, Y. J., and A. Di Gregorio. 2005. *Ciona intestinalis:* chordate development made simple. *Dev. Dyn.* **233:** 1–19.

Privat, A., C. Jacque, J. M. Bourre, P. Dupouey, and N. Baumann. 1979. Absence of the major dense line in myelin of the mutant mouse shiverer. *Neurosci. Lett.* **12:** 107–112.

Sanders, C. R., F. Ismail-Beigi, and M. W. McEnery. 2001. Mutations of peripheral myelin protein 22 result in defective trafficking through mechanisms which may be common to diseases involving tetraspan membrane proteins. *Biochemistry* **40:** 9453–9459.

Sasakura, Y., E. Shoguchi, N. Takatori, S. Wada, I. A. Meinertzhagen, Y. Satou, and N. Satoh. 2003. A genomewide survey of developmentally relevant genes in *Ciona intestinalis.* X. Genes for cell junctions and extracellular matrix. *Dev. Genes Evol.* **213:** 303–313.

Satoh, N. 2004. The ascidian tadpole larva: comparative molecular and development and genomics. *Nat. Rev. Genet.* **4:** 285–295.

Satou, Y., K. S. Imai, M. Levine, Y. Kohara, D. Rokhsar, and N. Satoh. 2003. A genomewide survey of developmentally relevant genes in *Ciona intestinalis.* I. Genes for bHLH transcription factors. *Dev. Genes Evol.* **213:** 213–221.

Schaeren-Wiemers, N., A. Bonnet, M. Erb, B. Erne, U. Bartsch, F. Kern, N. Mantei, D. Sherman, and U. Suter. 2004. The raft-associated protein MAL is required for maintenance of proper axon-glia interactions in the central nervous system. *J. Cell Biol.* **166:** 731–742.

Schafer, D. P., R. Bansal, K. L. Hedstrom, S. E. Pfeiffer, and M. N. Rasband. 2004. Does paranode formation and maintenance require partitioning of neurofascin 155 into lipid rafts? *J. Neurosci.* **24:** 3176–3185.

Schliess, F., and W. Stoffel. 1991. Evolution of the myelin integral membrane proteins of the central nervous system. *Biol. Chem. Hoppe-Seyler* **372:** 865–874.

Schneider, C., R. M. King, and L. Philipson. 1988. Genes specifically expressed at growth arrest of mammalian cells. *Cell* **54:** 787–793.

Schweitzer, J., T. Becker, C. G. Becker and M. Schachner. 2003. Expression of protein zero is increased in lesioned axon pathways in the central nervous system of adult zebrafish. *Glia* **41:** 301–317.

Shigetani, Y., F. Sugahara, Y. Kawakami, Y. Murakami, S. Hirano, and S. Kuratani. 2002. Heterotopic shift of epithelial-mesenchymal interactions in vertebrate jaw evolution. *Science* **296:** 1316–1319.

Sprinkle, T. J. 1989. 2′,3′-Cyclic nucleotide 3′-phosphodiesterase, an oligodendrocytes-Schwann cell and myelin-associated enzyme of the nervous system. *Crit. Rev. Neurobiol.* **4:** 235–301.

Stecca, B., C. M. Southwood, A. Gragerov, K. A., Kelley, V. L. Friedrich, Jr., and A. Gow. 2000. The evolution of lipophilin genes from invertebrates to tetrapods: DM-20 cannot replace proteolipid protein in CNS myelin. *J. Neurosci.* **20:** 4002–4010.

Suter, U. 2004. PMP22 gene. Pp. 547–564 in *Myelin Biology ansd*

*Disorders,* R. A. Lazzarini, J. W. Griffin, H. Lassmann, K.-A. Nave, R. H. Miller, and B. D. Trapp, eds., Elsevier Academic Press, Amsterdam.

**Swofford, D. L. 2003.** PAUP. *Phylogenetic Analysis Using Parsimony (and Other Methods).* Version 4. Sinauer Associates, Sunderland, MA.

**Taylor, V., and U. Suter. 1996.** Epithelial membrane protein-2 and epithelial membrane protein-3: two novel members of the peripheral myelin protein 22 gene family. *Gene* **175:** 115–120.

**Tsukita, S., and M. Furuse. 2002.** Claudin-based barrier in simple and stratified cellular sheets. *Curr. Opin. Cell Biol.* **14:** 531–536.

**Venkatesh, B., M. V. Erdmann, and S. Brenner. 2001.** Molecular synapomorphies resolve evolutionary relationships of extant jawed vertebrates. *PNAS* **98:** 11382–11387.

**Waehneldt, T. V., J.-M. Matthieu, and G. Jeserich. 1986.** Appearance of myelin proteins during vertebrate evolution. *Neurochem. Int.* **9:** 463–474.

**Werner, H., L. Dimou, M. Klugmann, S. Pfeiffer, and K. A. Nave. 2001.** Multiple splice isoforms of proteolipid M6B in neurons and oligodendrocytes. *Mol. Cell. Neurosci.* **18:** 593–605.

**Wilmot, G. R., P. A. Raymond, and B. W. Agranoff. 1993.** The expression of the protein p68/70 within the goldfish visual system suggests a role in both regeneration and neurogenesis. *J. Neurosci.* **13:** 387–401.

**Yan, Y., C. Lagenaur, and V. Narayanan. 1993.** Molecular cloning of M6: identification of a PLP/DM20 gene family. *Neuron* **11:** 423–431.

**Yoshida, M., and D. R. Colman. 1996.** Parallel evolution and coexpression of the proteolipid proteins and protein zero in vertebrate myelin. *Neuron* **16:** 1115–1126.

**Yoshida, M., W. S. Shan, and D. R. Colman. 1999.** Conserved and divergent expression patterns of the proteolipid protein gene family in the amphibian central nervous system. *J. Neurosci. Res.* **57:** 13–22.

**Zhou, Q., S. L. Wang, and D. J. Anderson. 2000.** Identification of a novel family of oligodendrocyte lineage-specific basic helix-loop-helix transcription factors. *Neuron* **25:** 331–343.

## Appendix

### Supplemental data on alignment of sea squirt and other sequences

Figure A1. Alignment of three sea squirt DM proteins with known sequences for human, chick, frog, zebrafish, spiny dogfish, and fruit fly. The amino acid positions are shown in parentheses just in front of each sequence and that of the overall alignment are shown at the top. The alignment is coded as follows: (1) White letters on a gray background represent blocks of identity. (2) Black letters on a gray background represent blocks with conserved replacements. (4) Black letters on a white background represent weak or no similarity. Sequences are as follows: human PLP (NP_963885), human DM20 (CAG38779), human M6A1 (NP 005268), human M6B (P35803), human M6B1 (NP 005268), human M6B2 (NP 001001996), human M6B3 (NP_005269), mouse DM20 (AAA39550), chick PLP (NP 990608), chick M6B (XP 416830), frog PLP (P35801), frog DMb (AAK17942), frog DMg1 (AAK17943), frog DMg2 (AAK17944), lungfish DM20 (BAB16472), spiny dogfish DMa (P36964), spiny dogfish DMb (I51324), spiny dogfish DMg (151325), zebrafish DM6a (NP 999852), zebrafish DMa2a (AAL59886), zebrafish PLP1a (NP 783166), zebrafish DMb1 (NP 998365), zebrafish DMg1 (NP 982251), fruit fly DM (AAF71284).

```
                              1                                                          60
Chick M6B             (1)   MYGVAADSHGKASPKGCSPTWAPLADAVRGLPSSSPPLPSPRLTAGKGHRGFSYSWGRWE

                              61                                                         120
Sea squirt DM-2       (1)   ----------MSVHSLHDLVEAESYQGYNIDRNGEVTEFYDHRNEDGTNSISTARHPLVT
Human M6B             (1)   -----------------------MKPAMETAAEENTEQSQERKVNSRAEMEIGRYHWMYPG
Human M6B1            (1)   -----------------------MKPAMETAAEENTEQSQER------------------
Human M6B2            (1)   -----------------------MKPAMETAAEENTEQSQER------------------
Human M6B3            (1)   -----------------------MKPAMETAAEENTEQSQER------------------
Chick M6B            (61)   RERRRGESACAGGRGAAFLVDGMKPAMETAAEENAEQSQEK------------------
Lungfish DM20         (1)   ---------------------------------------------------------M
Zebrafish DMg1        (1)   -----------------MDGMKP--ESTAEENQDERDES-------------------

                              121                                                       180
Sea squirt DM-1       (1)   ---------------------MGCFDSCLRCLS---LVPWASLIAAVLCWAGTALFCGTA
Sea squirt DM-2      (51)   TSHPATTVATTPKRKKVRYRDAGCDGCVNCLITIGTMPFATAFAVLLCWLGTAMFIGGG
Human DM20            (1)   --------------------MGLLECCARCLV---GAPFASLVATGLCFFGVALFCGCG
Human PLP            (1)   --------------------MGLLECCARCLV---GAPFASLVATGLCFFGVALFCGCG
Human M6A1           (1)   ----------MEENMEEGQTQKGCFECCIKCLG---GIPYASLIATILLYAGVALFCGCG
Human M6B           (39)   SKNHQYRPVPNLGDRAGPLSSPGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCG
Human M6B1          (20)   --------------------KGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCG
Human M6B2          (20)   --------------------KGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCG
Human M6B3          (20)   --------------------KGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCG
Mouse DM20           (1)   --------------------MGLLECCARCLV---GAPFASLVATGLCFFGVALFCGCG
Chick PLP            (1)   --------------------MGLLECCARCLI---GAPFASLVATGLCFFGVALFCGCG
Chick M6B          (102)   --------------------KGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCG
Frog PLP             (1)   --------------------MGWHDGCIRCMV---GVPFASVIATVLCFAGVALFCGCG
Frog DMb             (1)   ----------MEENMEEGQTQKGCFECCIKCLG---GIPYASLIATILLYAGVALFCGCG
Frog DMg1            (1)   --------------------MGCFECCIKCLG---GVPYASLVATILCFCGVALFCGCG
Frog DMg2            (1)   --------------------MGCFECCIKCLG---GVPYASLVATILCFSGVALFCGCG
Lungfish DM20        (2)   GKRKVDSAQGRFPPGERSGLFYGCVECCAKCLG---GIPYASMVATILCFSGVALFCGCG
Zebrafish DM6a       (1)   --------------------MGCSECCLKCLS---GIPYASLIATILLYAGVALFCGCG
Zebrafish DMa1       (1)   --------------------MGCYECCMRCLG---GVPYVSLAATLCFSGIALFCGCG
Zebrafish PLP1a      (1)   --------------------MGCYECCMRCLG---GVPYVSLAATLCFSGIALFCGCG
Zebrafish DM6b1      (1)   ----------MEEDMDEGQTQKGCMECCLKCLG---GIPYPSLIATILLYAGVALFCGCG
Zebrafish DMg1      (21)   --------------------KGCFECCIKCLG---GVPYASLIATILCFSGVALFCGCG
Spiny dogfish DMg    (1)   --------------------MGCFECCIKCLG---GVPYASLIATILCFSGVALFCGCG
Spiny dogfish DMa    (1)   --------------------MGCSECCVRCLG---GVPYASLIATILCFVGVALFCGCG
Spiny dogfish DMb    (1)   ----------MEENMEEGQTQKGCFECCIKCLG---GIPYASLIATILLYAGVALFCGCG
Fruit fly DM         (1)   --------------------MGECCQSCMAR---IPYATLIATLMCLLGVGIFCFTM

                              181                                                       240
Sea squirt DM-1      (37)   HEAVSNSYILLNSTDFYWNET----------------------------STDTIIQNEK
Sea squirt DM-2     (111)   WSALDGTMNLWAQFGPMAGNPYNHPPPQYLYTGNALTPQFINLNRDDLGNRVPAVFQAIK
Sea squirt DM-3       (1)   -----------MNYNQRGN------------------------------LVPQVFAGIQ
Human DM20           (37)   HEALTGTEKLIETYFSKNYQ--DYE-----------------------YLINVIHAFQ
Human PLP            (37)   HEALTGTEKLIETYFSKNYQ--DYE-----------------------YLINVIHAFQ
Human M6A1           (48)   HEALSGTVNILQTYFEMARTAGDTL-----------------------DVFTMIDIEK
Human M6B            (96)   HVALAGTVAILEQHFSTNTS--DHA-----------------------LLSEVIQLMQ
Human M6B1           (56)   HVALAGTVAILEQHFSTNAS--DHA-----------------------LLSEVIQLMQ
Human M6B2           (56)   HVALAGTVAILEQHFSTNAS--DHA-----------------------LLSEVIQLMQ
Human M6B3           (56)   HVALAGTVAILEQHFSTNAS--DHA-----------------------LLSEVIQLMQ
Mouse DM20           (37)   HEALTGTEKLIETYFSKNYQ--DYE-----------------------YLINVIHAFQ
Chick PLP            (37)   HEALTGTEQLIETYFSKNYQ--DYE-----------------------YLIDVIHAFQ
Chick M6B           (138)   HVALTGTVSILEQHFSTTAS--DHA-----------------------LLSEVIQLMQ
Frog PLP             (37)   HEALSGTEKLIETYFSKNYQ--EYE-----------------------YLIHVINAFQ
Frog DMb             (48)   HEALTGTVNILQTYFEMARTAGDTL-----------------------DVFTMIDIEK
Frog DMg1            (37)   HVALTGTVTILEQHFSSNPS--DHV-----------------------LLTEVIQVMQ
Frog DMg2            (37)   HVALTGTVAILEQHFSSNPS--DHV-----------------------FLTEVIQLMQ
Lungfish DM20        (59)   HEALNGTEYLMETHFIRNYQ--HFQ-----------------------ALSNLIDAFQ
Zebrafish DM6a       (37)   HEALSGTVTILQTYFDVVRSPVDAL-----------------------DVFTMIDIEK
Zebrafish DMa1       (37)   HQALTESERLIETYFARDNQ--DYY-----------------------TLVYLIEYFQ
Zebrafish PLP1a      (37)   HQALTESERLIETYFARDNQ--DYY-----------------------TLVYLIEYFQ
Zebrafish DM6b1      (48)   HEALSGTVTILQNYFEVIRGPGDGL-----------------------DVFTIIDIIK
Zebrafish DMg1       (57)   HVALTGTVTMVENHFSRISS--DHA-----------------------ALTDVVQLLQ
Spiny dogfish DMa    (37)   HEALTGTEKLIELYFSNDFM--DYA-----------------------LLVNVIQVFQ
Spiny dogfish DMb    (48)   HEALSGTVTILQNNFEVVRGAGDTI-----------------------DVFTMIDIEK
Spiny dogfish DMg    (37)   HVALTKVERIVQLYFSNNAS--DHV-----------------------LLTDVIQMMH
Fruit fly DM         (35)   YRGASLTVIMVDQVFHLRLI----W-----------------------IEAVQ
```

```
                                     241                                               300
Sea squirt DM-1    (68)  YSIYGIAAFMFVLSILLFVDGILSTRVVKSDYEEGCKTSCCGLCIGIFYTM---------
Sea squirt DM-2   (171)  YSFYGLIPEMFVFTIMLACDNRHSSRAMASKKRS-CKSSSSGICLTAALIG---------
Sea squirt DM-3    (19)  YYLIALGVEMLINCTVFMVDSVRSKKAVKDRTYQGKRTGLSRISCTGWLAS---------
Human PLP          (70)  YVIYGTASFFFLYGALLLAEGFYTTGAVRQIFGD-YKTTICGKGLSATVTGGQ-KGRGSR
Human DM20         (70)  YVIYGTASFFFLYGALLLAEGFYTTGAVRQIFGD-YKTTICGKGLSATFVG---------
Human M6A1         (83)  YVIYGIAAFVVYGILLMVEGFTTGAIKDLYGD-FKITTCGRCVSAWFIM---------
Human M6B         (129)  YVIYGIASFFFLYGIILLAEGFYTTSAVKELHGE-FKTTACGRCISGMFVF---------
Human M6B1         (89)  YVIYGIASFFFLYGIILLAEGFYTTSAVKELHGE-FKTTACGRCISGMFVF---------
Human M6B2         (89)  YVIYGIASFFFLYGIILLAEGFYTTSAVKELHGE-FKTTACGRCISGMFVF---------
Human M6B3         (89)  YVIYGIASFFFLYGIILLAEGFYTTSAVKELHGE-FKTTACGRCISGMFVF---------
Mouse DM20         (70)  CVIYGTASFFFLYGALLLAEGFYTTGAVRQIFGD-YKTTICGKGLSATFVG---------
Chick PLP          (70)  YVIYGTASFFFLYGALLLAEGFYTTGAVRQIFGD-YRTTICGKGLSATVTGGP-KGRGAR
Chick M6B         (171)  YVIYGIASFFFLYGIILLAEGFYTTSAVKELHGE-FKTTACGRCISGMFVF---------
Frog PLP           (70)  YVIYGIAIFFFLFGILLLAEGFYTTAIKHILGE-FKPPAIKGGLISTVTGGTPKGRSTR
Frog DMb           (83)  YVIYGIAAFFVYGILLMVEGFTTGAIKDLYGD-FKITTCGRCVSAWFIM---------
Frog DMg2          (70)  YVIYGIASFFFLYGIILLAEGFYTTSAVKELHSE-FKTTACGRCISGMFVF---------
Lungfish DM20      (92)  YVIYGMASFFFLYGVLLLAEGFYTTSSVIKKVFGE-NKTTLCGRCISASFIG---------
Zebrafish DM6a     (72)  YVIYGVAAFFVYGILLMVEGFTTGAIRDLYGD-FKITTCGRCVSAWFIM---------
Zebrafish DMa1     (70)  YIIYGLASFFFLYCIVLLAEGFYTTSAVRQSLGE-FRSTVCGRCLSTTFIV---------
Zebrafish PLP1a    (70)  YIIYGLASFFFLYCIVLLAEGFYTTSAVRQSLGE-FRSTVCGRCLSTTFIV---------
Zebrafish DM6b1    (83)  YVIYGIASAFFVYGILLMVEGFFTSGAIKDLYGD-FKITTCGRCVSAWFIM---------
Zebrafish DMg1     (90)  YIIYGIACLFFLYGIILLAEGFYTTSAVKEMHSE-FKTTVCGRCISAMFVF---------
Spiny dogfish DMa  (70)  YIIYGTASFSFLYGVLLLAEGFYTTSAVKALFGE-FRTTVCGRCVSATFIF---------
Spiny dogfish DMb  (83)  YVIYGVAAFFVYGILLMVEGFTTGAIKDLYGD-FKITTCGRCVSGWFIM---------
Spiny dogfish DMg  (70)  YVIYGVASFSFLYGIILLAEGFYTTSAVKEIHGE-FKTTVCGRCISGMSVF---------
Fruit fly DM       (61)  MIFVILGAGMAALGFMILFVGFLATGATRYKVYRAWRSRVGGRISCAVLMG---------


                                     301                                               360
Sea squirt DM-1   (119)  -------------------------LVYICGVGWLLITCFAAFPIHFFIVANSQCKAN
Sea squirt DM-2   (221)  -------------------------VTYILILAWVFFLAFTTLGVYYYRISMLRCSDL
Sea squirt DM-3    (70)  -------------------------YSAFLALMWFVMACFGVIPLLEFRITMQRCWDL
Human PLP         (128)  GQHQAHSLERVCHCLGKWLGHPDKFVGITYALTVVWLLVFACSAVPVYIYFNTWTTCQSI
Human DM20        (120)  -------------------------ITYALTVVWLLVFACSAVPVYIYFNTWTTCQSI
Human M6A1        (133)  -------------------------LTYLFMLAWLGVTAFTSLPVYMYFNLWTTICRNT
Human M6B         (179)  -------------------------LTYVLGVAWLGVFGFSAVPVFMFYNIWSTCEVT
Human M6B1        (139)  -------------------------LTYVLGVAWLGVFGFSAVPVFMFYNIWSTCEVI
Human M6B3        (139)  -------------------------LTYVLGVAWLGVFGFSAVPVFMFYNIWSTCEVI
Mouse DM20        (120)  -------------------------ITYALTVVWLLVFACSAVPVYIYFNTWTTCQSI
Chick PLP         (128)  GPQRAHSLQRVCQCLGKWLGHPDKFVGITYVLTIVWLLAFACSAVPVYIYFNTWTTCQSI
Chick M6B         (221)  -------------------------LTYVLGVAWLGVFGFSAVPVFMFYNIWSTCEVI
Frog PLP          (129)  GRQPVHTIELICRCLGKWLGHPDKFVGVTYIITILWILLIFACSAVPVYIYFNTWVTCQSI
Frog DMb          (133)  -------------------------LTYLFMLAWLGVTAFTSLPVYMYFNWTTICRNA
Frog DMg1         (120)  -------------------------LTYILGVAWLGVFGFSAIPVFMFYNMWSSCEFI
Frog DMg2         (120)  -------------------------LTYILGVAWLGVFGFSAIPVFMFYNMWSSCEVI
Lungfish DM20     (142)  -------------------------VTYALVLIWLAIFVFTALPLYIFHNMWSTCKAI
Zebrafish DM6a    (122)  -------------------------LTYIFMLAWLGVTAFTSLPVFMYFNIWNTCQNM
Zebrafish DMa1    (120)  -------------------------ITYFLVVVWLLVFSLSAFPVYFFYNMAKTCQTI
Zebrafish PLP1a   (120)  -------------------------ITYFLVVVWLLVFSLSAFPVYFFYNMAKTCQTI
Zebrafish DM6b1   (133)  -------------------------LTYIFMLAWLGVTAFTSLPVFIYFNIWIICQNT
Zebrafish DMg1    (140)  -------------------------LTYILGLGWLGIFGFSAVPVFLFYNMWTTCGAM
Spiny dogfish DMa (120)  -------------------------LTYALGVTWMGVFAFSALPVYIYYTMWSTCQMV
Spiny dogfish DMb (133)  -------------------------LTYIFMLAWLGVTAFTSLPVFMYFNIWILCQNV
Spiny dogfish DMg (120)  -------------------------LTYLIGIAWLGVFGFSAVPAFIYYNMWSACQTI
Fruit fly DM      (112)  -------------------------ITYLLNFVWSLILCFLVVVTFIYTMFWNMCTSV
```

```
                        361                                                        420
Sea squirt DM-1   (152) --------------TCIDFRQTGILPAHLPDDI--GRICG---DDLTKFCSTP----TFN
Sea squirt DM-2   (254) Q--EREFSTGVLKDVCIDLVQLGVMMFKNTQDNGFGKLCGPG-DSMQSYGKLEGYCQNYY
Sea squirt DM-3   (103) ANPNWS--PGTAQRICIDLVQYGLVFFKETQDTGYGLICGPSNNLLYPDGNLEGFCANYY
Human PLP         (188) AFPSKT--SASIGSLCADARMYGVLPWNAFP----GKVCG---SNLLSICKTA----EFQ
Human DM20        (153) AFPSKT--SASIGSLCADARMYGVLPWNAFP----GKVCG---SNLLSICKTA----EFQ
Human M6B         (212) KSPQSNG-TSGVEQICVDVRQYGIIPWNAFP----GKICG---SALENICNTN----EFY
Human M6A1        (166) TLVEG-------ANLCIDLRQFGIVTIGEEK-----KICTVS-ENFLRMCEST----ELN
Human M6B1        (172) KSPQTNG-TTGVEQICVDIRQYGIIPWNAFP----GKICG---SALENICNTN----EFY
Human M6B2        (172) KSPQTNG-TTGVEQICVDIRQYGIIPWNAFP----GKICG---SALENICNTN----EFY
Human M6B3        (172) KSPQTNG-TTGVEQICVDIRQYGIIPWNAFP----GKICG---SALENICNTN----EFY
Mouse DM20        (153) AFPSKT--SASIGSLCADARMYGVLPWNAFP----GKVCG---SNLLSICKTA----EFQ
Chick M6B         (254) KSLQINM-TVPGDQICVDIRQYGIIPWNAVP----GKACG---PILENICNTN----EFY
Frog PLP          (189) AFPGKT--TTSVSTLCSDRRMYGVLPWNAFP----GKVCG---TSLLAICKTS----EFQ
Frog DMb          (166) TLIEE-------ANFCIDLRQFGIVTVGEEK-----KLCSPS-DNFLRMCEST----ELN
Frog DMg1         (153) KILPTNL-TTTADQICVDIRQFGIIPWNALP----GKACG---QALEQICNSN----EFY
Frog DMg2         (153) KALPTNL-TTTADQICVDIRQFGIIPWNALP----GKACG---QALEQICNSN----EFY
Lungfish DM20     (175) ELP-EN--SGRIHQLCVDARQYGILPWSATP----GEVCG---YNLQTVCRTR----EFG
Zebrafish DM6a    (155) TLLEGSLSSANSANLCFDLRQYGIVSIGEEK-----RLCADN-ENFKKMCESN----ELD
Zebrafish DMa1    (153) NFLSET--PSSINQLCIDARQYGLLPWTAIP----GKACG---MTLSNVCKTR----EFS
Zebrafish PLP1a   (153) NFLSET--PSSINQLCIDARQYGLLPWTAIP----GKACG---MTLSNVCKTR----EFS
Zebrafish DM6b1   (166) TVLQG-------ASLCIDPRQFGVVPIGEEK-----TLCVGS-ENFFKMCESN----ELD
Zebrafish DMg1    (173) KSPIAN--LTSVDNICVDVRQYGIIPWNANP----GKACG---STLSDICNTS----EFY
Spiny dogfish DMa (153) KYVTEN--GTGFDDICVDARQYGILPWNASP----GKICG---LSLAAVCNTS----EFE
Spiny dogfish DMb (166) TIMES-------TDLCFDLRQFGIVPIHEQK-----TVCTLN-ENFSKLCQSN----DLN
Spiny dogfish DMg (153) SSPPVNL-TTVIEEICVDRQYGIIPWNASP----GKACG---STLTTICNTS----EFD
Fruit fly DM      (145) EHSQS---------CIDLTQFHFMFPPNTKLED-MKVCEK--YEIKAFCKDG----VEN


                        421                                                        480
Sea squirt DM-1   (189) VTYQLYIVAFAGAAITLLSMKQFLMCLSANFAYLKMTKKLAVYEDTKYREEME-------
Sea squirt DM-2   (311) VVYQYMLVAFSGCLLNIAMINFIMILTINYNDLARKRAAPTRSTASAAPSQA-------
Sea squirt DM-3   (161) PVFTLYVVSFAGAAITMMSMILFALRHASNFVYLKVRRWENPLGPPGGKKAPKGPHPHGY
Human PLP         (235) MTFHLFIAAFVGAAATLVSLLTFMIAATYNFAVLKLMGRG-TKF----------------
Human DM20        (200) MTFHLFIAAFVGAAATLVSLLTFMIAATYNFAVLKLMGRG-TKF----------------
Human M6A1        (209) MTFHLFIVALAGAGAVIAMVHYLMVLSANVAYVKDACRMQKYEDIKSKEEQE-------
Human M6B         (260) MSYHLFIVACAGAGATVIALIHFLMILSSNWAYLKDASKM-QAYQDIKAKEEQ-------
Human M6B1        (220) MSYHLFIVACAGAGATVIALLIYMMATTYNYAVLKFKSRE-DCCTKF-------------
Human M6B2        (220) MSYHLFIVACAGAGATVIALLIYMMATTYNYAVLKFKSRE-DCCTKF-------------
Human M6B3        (220) MSYHLFIVACAGAGATVIALLIYMMATTYNYAVLKFKSRE-DCCTKF-------------
Mouse DM20        (200) MTFHLFIAAFVGAAATLVSLLTFMIAATYNFAVLKLMGRG-TKF----------------
Chick PLP         (235) MTFHLFIAAFVGAAATLVSLLTFMIAATYNFAVLKLMGRG-TKF----------------
Chick M6B         (302) MSYHLFIVACAGAGATVIALFLMILSSNWAYLKDASKM-QAYQDIKAKEEQ-------
Frog PLP          (236) MTFHLFIAAFVGAAATLVSLLTYMVGASENVAVLKLRVTGRS-DRSKF-----------
Frog DMb          (209) MTFHLFIVALAGAGAAVIAMVHYLMVLSANWAYVKDACRMQKYEDIKSKEEQE-------
Frog DMg1         (201) MSYHLFIVACAGAGATVIALIHFLMILSSNWAYLKDASKM-QVYQDIKAKEEQ-------
Frog DMg2         (201) MSYHLFIVACAGAGATVIALLIYMMATTYNYAVLKFKSRE-DCCTKF-------------
Lungfish DM20     (221) MTYHLFIATFVGAGATAVALLTYMMSTTYNFAIVKILGRE-DYCTKF-------------
Zebrafish DM6a    (205) LTFHLFVCALAGAGAAVIAMVVAVSVLINNHASLTSKNAGRYCTRF--------------
Zebrafish DMa1    (200) LTYNLYIAAFAGAGMTVLALLTYTVSTTYNFAVLRYLGRK-GISPRC------------
Zebrafish PLP1a   (200) LTYNLYIAAFAGAGMTVLALLTYTVSTTYNFAVLRYLGRK-GISPRC------------
Zebrafish DM6b1   (209) MTFHLFVCALAGAGAAVIAMIHYLMVLSANWAYVKDACKMQKYEDCKSKEEQE-------
Zebrafish DMg1    (220) LSYHLYIVACAGAGATVIALIHYLMILAANWAYLKGACQQTHAYQDIKTRDDQ-------
Spiny dogfish DMa (200) LTYHLFIATFAGAAATVIALLTYMMSSTYNYAVLKFLSRD-DCCTKF-------------
Spiny dogfish DMb (209) MTFHLFIVALAGAGAAVIAMVHYLMVLSANWAYVKDACRMQKYEDIKSKEEQE-------
Spiny dogfish DMg (201) LSYHLFIVACAGAGATVIALLIYMMATTYNFAVLKFKSRE-DCCTKF-------------
Fruit fly DM      (188) AEVMFILATLLSTLLVLLSLVHYLMCLSANYAHIRDHEKFQELQEIQNLNELE-------


                        481                                                        540
Sea squirt DM-1   (242) LNDIINTARSN--------------------ERLTYKY---------------------
Sea squirt DM-2   (364) STTPIVARSVR--------------------SIPNTSVGPPPYVVSNRDYDRERRSSKDI
Sea squirt DM3    (221) DDRPYTTNSLQNPNYEHNVLLVDPDRMPDDISLNPYSLQPGDDVMDMKQRYFDRADGDHR
Human M6A1        (262) LHDIHSTRSKE--------------------RLNAYT---------------------
Human M6B         (312) ELQDIQSRSKE--------------------QLNSYT---------------------
Chick M6B         (354) ELQDIQSRSKE--------------------QLNSYT---------------------
Chick PLP         (278) ---------------------------------------------------------
Frog DMb          (262) LHDIHSTRSKE--------------------RLNAYT---------------------
Frog DMg1         (253) ELQDIQSRSKK--------------------QINSYT---------------------
Zebrafish DM6b1   (262) LHDIHSTRSKE--------------------RLNAYT---------------------
Zebrafish DMg1    (273) ELQDVQSRSKE--------------------GLNSYS---------------------
Spiny dogfish DMb (262) LHDIHSTRSKE--------------------RLNAYT---------------------
Fruit fly DM      (241) YSATSKDRF------------------------------------------------
```

```
                         541                                                      600
Sea squirt DM-2    (404) ELDYMLDSASSYERPREEDSHMGVNYYYRR----------------------------
Sea squirt DM-3    (281) AYDKRRSSRSSSSYADEKPLTYDSKRRSRRRSHDSRETPRRRRSDVSDESYDHGSRHPTS


                         601                                                      660
Sea squirt DM-3    (341) RSHRGSKSRDLSTASENGDVMKGTGSSDEYERSSAPYRPKPLRRKSKGAKSNPEIYKQSS


                         661                                     701
Sea squirt DM-3    (401) RRRKRSSHRRDESDSPPSDYYLQEDLKRRSSTFNDRSRSVV
```