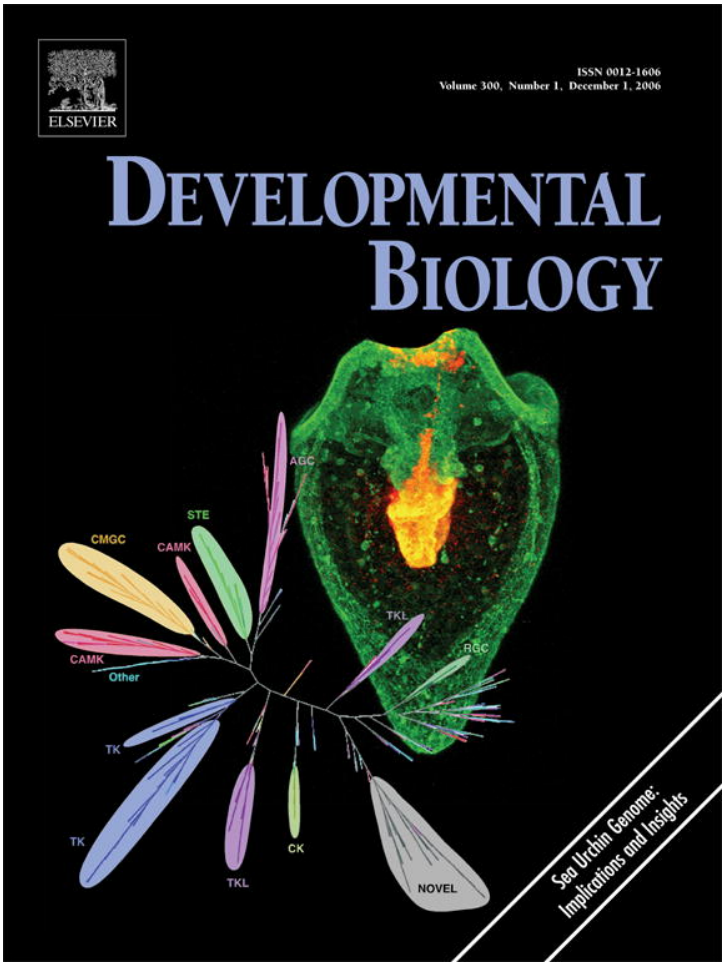


Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 300 (2006) 267–281
DEVELOPMENTAL BIOLOGY
www.elsevier.com/locate/ydbio

Sea urchin metalloproteases: A genomic survey of the BMP-1/tolloid-like, MMP and ADAM families

Lynne Angerer [a,*], Sofia Hussain [b], Zheng Wei [a], Brian T. Livingston [b]

[a] National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD 20815, USA
[b] Department of Biology, University of South Florida, SCA110, Tampa, FL 33620, USA

Received for publication 15 May 2006; revised 25 July 2006; accepted 26 July 2006
Available online 12 August 2006

Abstract

Analysis of the *Strongylocentrotus purpuratus* genome has revealed approximately 240 metalloprotease genes, and they represent all 23 families expressed in vertebrates. EST/cDNA sequencing and microarray analysis show that nearly 70% are represented in embryo RNA. Among them are many metalloproteases with demonstrated developmental roles in other systems-BMP-1/TLD (tolloid) (astacins), MMPs (matrix metalloproteases) and the ADAMs (disintegrin/metalloproteases). The developmental functions of these kinds of metalloproteases include modifying the extracellular matrix, regulating signaling pathways or modulating cellular adhesive properties. The unexpectedly large number of BMP-1/TLD-like protease genes (23) results primarily from expansion of a set encoding an unusual domain conserved in structure and primary sequence only in nematode astacins. Such proteases may have interesting developmental functions because the expression patterns of several are highly regulated along the primary axis at times when cell differentiation and morphogenesis begin. The size of the sea urchin MMP family and the clustered arrangement of many of its members are similar to vertebrates, but phylogenetic analyses suggest that different ancestral genes were independently amplified in sea urchins and vertebrates. One expansion appears to be genes encoding MMPs that have putative transmembrane domains and may be membrane-tethered (MT). Interestingly, the genes encoding TIMPs, inhibitors of MMPs, have also been amplified and the 10 genes are tandemly arranged in a single cluster. In contrast, there are fewer ADAM and ADAMTS genes in sea urchins, but they represent all but one of the chordate-specific groups. The genome sequence now opens the door to experimental manipulations designed to understand how modulation of the extracellular environment affects development.

Keywords: Annotation; Microarray; Evolution; Gene expression; Embryo

Introduction

Proteolytic processing of extracellular matrix and proteins tethered to the cell membrane play a critical role in tissue remodeling, cell migration, cell differentiation and morphogenesis during embryonic development (for reviews, see Sarras, 1996; Vu and Werb, 2000; Blobel, 2005). In particular, these activities are critical for cell movements and for regulating the concentrations of signaling molecules such as growth factors and cytokines. Many extracellular proteolytic activities are carried out by members of the superfamily of proteins called metzincins, which share a zinc atom in the catalytic domain. Within this superfamily are the astacin/BMP-1/tolloid metalloproteases, the matrix metalloproteinases (MMPs) and A Disintegrin and Metalloproteases (ADAMS). These groups of proteases are known to be required for normal cellular and developmental processes. For example, tolloid-mediated proteolysis of chordin promotes BMP signaling (reviewed by Mullins, 1998); MMPs can degrade most of the components of the ECM, thereby regulating cell migration, apoptosis, as well as releasing and activating growth factors (reviewed by Vu and Werb, 2000); and ADAMs function to release membrane-anchored growth factors, cytokines and receptors (reviewed by Blobel, 2005).

In the sea urchin embryo, many different signaling pathways have been shown to be required for cell fate specification and morphogenesis. Canonical Wnt signaling is necessary and

* Corresponding author.
E-mail address: langerer@mail.nih.gov (L. Angerer).

0012-1606/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2006.07.046

sufficient for endomesoderm specification (Logan et al., 1999; Wikramanayake et al. 1998; Emily-Fenouil et al., 1998; Vonica et al., 2000); Nodal (Duboc et al., 2004, 2005) and BMP (Angerer et al., 2000; Duboc et al., 2004) signaling regulate oral and aboral ectoderm development; Notch signaling is required for specification of mesenchymal cell types (Sherwood and McClay, 1999; Sweet et al., 1999) and regulating the ectoderm/endoderm boundary (Sherwood and McClay, 2001); and VEGF signals are required for skeletogenic mesenchyme differentiation (C. Gache, personal communication). Furthermore, there is clear evidence that the extracellular matrix environment is important for normal morphogenesis of sea urchin embryos. Immune interference with either the fibropellins (Burke et al., 1991) or the hyalin layer in the apical ECM (Adelson and Humphreys, 1988) inhibits normal gastrulation as does blocking collagen assembly within the blastocoel (Wessel and McClay, 1987). Of particular relevance to this paper is the observation that inhibitors of the tolloid-like metalloprotease, procollagen C-terminal proteinase, block gastrulation and spiculogenesis (Huggins and Lennarz, 2001). In other systems, MMP activities are associated with the migratory and invasive properties of cells. In the sea urchin, there are several cell types displaying such properties. For example, primary mesenchyme cells ingress through the basement membrane and migrate to specific locations in the blastocoel and other mesenchyme cells invade the ectoderm during gastrulation.

Only six genes encoding metalloproteases in these three metzincin families have been examined in the sea urchin embryo. Within the astacin (BMP-1/TLD) family, *SpAN* (Reynolds et al., 1992) (*BP10* in *Paracentrotus lividus*; Lepage et al., 1992a) is transiently expressed in non-vegetal blastomeres during early blastula stages. SpAN protein accumulates in hyalin-containing extracellular matrix and can cleave hyalin (our unpublished observations), raising the possibility that during development it may remodel this substrate, which is required for normal cell movements associated with gastrulation (Adelson and Humphreys, 1988). A protein closely related to human BMP-1 (suBMP-1) is secreted by most cells of the gastrula-stage embryo and accumulates in both apical and basal extracellular matrices (Hwang et al., 1994). As mentioned above, the tolloid-like metalloprotease, suBMP-1, may function in spicule elongation and gastrulation (Huggins and Lennarz, 2001), both of which depend on the apical extracellular matrix (Zito et al., 2000, 2003; and references cited above). Spiculogenesis also requires MMP activity, as shown by the use of specific inhibitors (Ingersoll and Wilt, 1998). Several MMP candidates have been identified that are membrane-tethered, and these have interesting spatially restricted expression patterns (Ingersoll and Pendharkar, 2005). Another sea urchin *MMP* gene encodes the hatching enzyme (Ghiglione et al., 1994), which is expressed in the same pattern as *SpAN* (Reynolds et al., 1992; Lepage et al., 1992b). Only one gene encoding an ADAM has been described; *SpADAM* is most closely related to ADAMs 12, 13 and 19, and its early uniform expression pattern is progressively restricted to mesenchyme and neuronal cells during morphogenesis (Rise and Burke, 2002).

Searches of the genome sequence have uncovered a large set of genes, approaching 240, in the superfamily encoding metalloproteases. The BMP-1/Tld, MMP and ADAMs families are also large. The majority of genes in the superfamily, as well as in these three families, are expressed during embryogenesis. Nearly all encode proteases linked to a diverse set of other domains. In order to compare the sea urchin metalloproteases to those found in other organisms and to gain insight into the evolution of this large group of proteins, we have taken two approaches. First, we compared the collection of domain architectures within each group to that in other organisms and second, we employed phylogenetic methods to further assess relationships among genes. The most interesting findings are: (1) There are 11 genes encoding astacin proteases with an unusual domain structure conserved only in nematodes, and we show that one of them is expressed in a striking and novel pattern in the embryo; (2) the relatively large set of MMP genes is derived largely from duplications of ancestral genes different from those that generated the vertebrate set; (3) there has been an expansion of TIMP genes, and all 10 are tandemly arranged in a single cluster; and (4) while the ADAMs gene family is significantly smaller in sea urchins compared to vertebrates, it contains representatives of all the chordate-specific types.

Materials and methods

Methods are provided in the figure legends.

Results and discussion

Identification of the S. purpuratus metalloprotease gene superfamily

We used the protease domains of metalloprotease genes found in metazoan genomes to query the sea urchin genome (the set of glean3 gene predictions 2005 07 18 using the 6× WGS) and identified 233 sequences that represent 23 families of metalloproteases (http://merops.sanger.ac.uk/). These are listed in Supplementary Table 1. Most of the vertebrate families are represented in the sea urchin genome at about the same frequency; the three groups whose numbers differ in the sea urchin and human genomes by greater than a factor of 2 are highlighted with gray shading in Table 2 (supplemental data). It is important to note that while all of the sea urchin gene predictions contain sequence encoding the protease domains, some are incomplete, either because of prediction errors or because portions of genes are found on different scaffolds in the assembly that was queried. There are several sources of error that affect the accuracy of gene numbers. Some genes may not be included in the glean3 prediction set. For example, searches of the whole genome and the glean3 set for one family of proteases revealed that the glean3 set lacked 1 gene out of 24 or 4%. On the other hand, there may be few cases where haplotypes or duplicates resulting from assembly problems remain, although we removed many of these using ClustalW-based sequence alignment methods. Therefore, while small

errors exist, the final tally is likely to be reasonably accurate, because these errors are at least partially offsetting.

The microarray data provided by Samanta et al. (in press) shows that a surprisingly high fraction (70%) of the sequences are represented in sea urchin embryo RNA (between egg and gastrula stage). Genes were scored as transcribed if most of the predicted exons and the 3′ UTR regions showed signals greater than 3. This criterion selects against false positives resulting from cross hybridization to patches of conserved sequence. The expression survey results for each metalloprotease family are presented in Fig. 1. Shaded in gray are the three families that we will focus on here because they contain members known to have important developmental functions in other systems (BMP-TLD, MMPs and ADAMs).

BMP-1/Tolloid (astacin; M12A metalloproteases)

The astacin proteases are a subclass of metzincins that are widespread among different animal phyla (reviewed in Mohrlen et al., 2003). These very ancient proteases are also found in bacteria and fungi, in which the vast majority of representatives consist only of the protease domain, the one known exception having a linked ricin domain. During the evolution of the metazoan clades, a variety of different domains have accumulated on the C-terminal side of the protease domain. These include the ShKToxin and MAM domains in the Cnidaria. In both protostome and deuterostome lines, some astacin proteases appeared that lacked the cnidarian features while other groups acquired the protein–protein interaction domains, thrombospondin, CUB and EGF-like, in diverse combinations. Finally, within the vertebrates yet another architecture emerged in the meprins, which lack TSP1 and CUB domains but contain the MAM/MATH combination, sometimes associated with a transmembrane domain that tethers these extracellular proteases to the plasma membrane.

The sea urchin genome contains substantially more astacin genes than other organisms. Only nematode genomes contain more of these genes (Mohrlen et al., 2003). In the sea urchin, these vary in structure from those containing only the protease domain, to those containing ShKT domains to those with a large number of combinations of CUB and EGF-like domains, including the most complex forms that are characteristic of the tolloid (TLD)/BMP-1 proteases (reviewed in Sarras, 1996; Mohrlen et al., 2006). The known functions of these proteases and the variable domain structure suggest that this is a versatile protease recruited for diverse functions via protein–protein interactions.

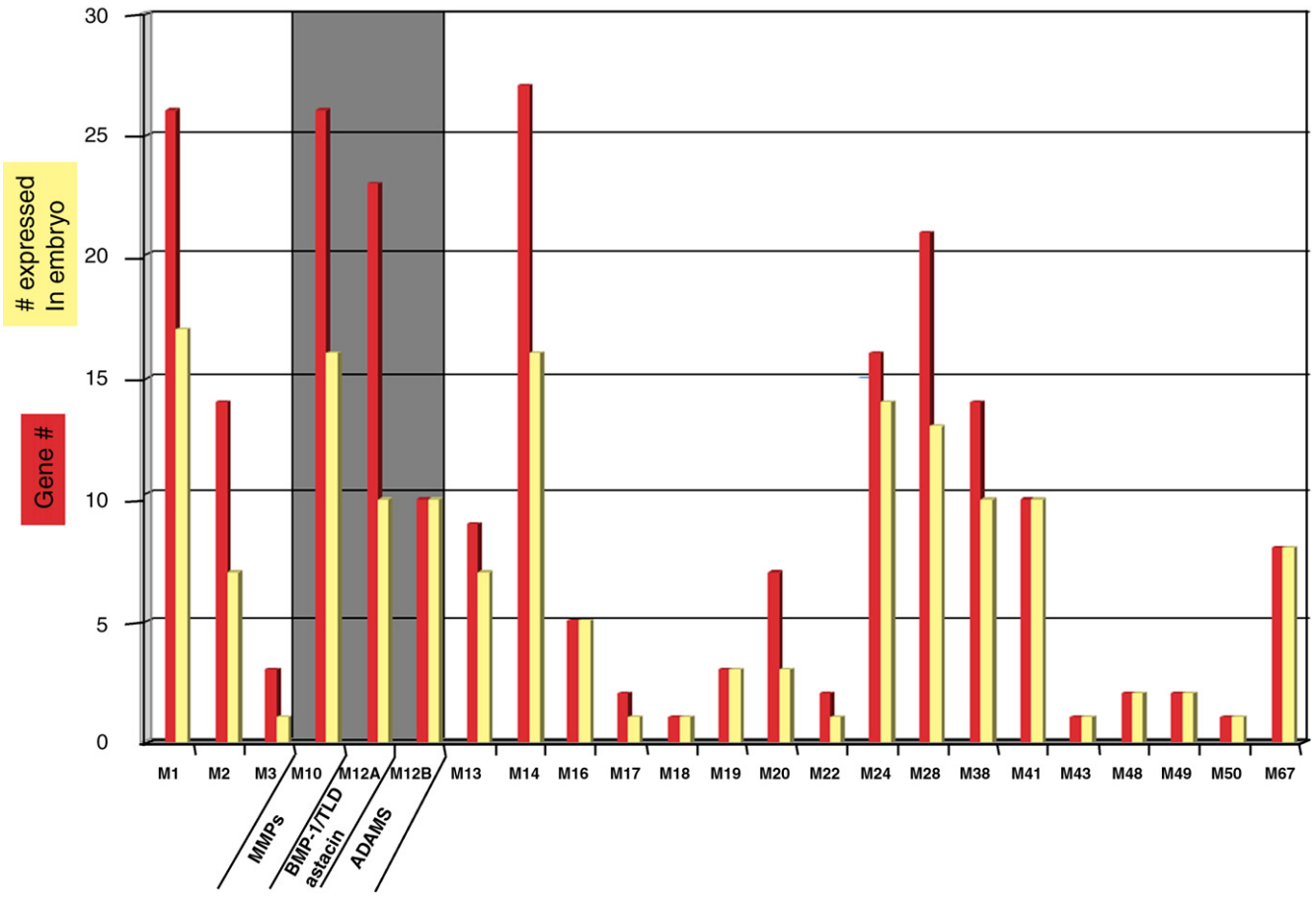


Fig. 1. The representation of sea urchin genes in different families within the metalloprotease family. Genes were identified using human metalloprotease domains as queries in BLAST searches. Gene identities were further verified by reciprocal BLAST analysis. Duplicates and haplotypes were removed after comparing ClustalW sequence alignments. Genes were scored as expressed during the interval from egg to gastrula stage, using the microarray data of Samanta et al. (in press). See text for details.

The different structural types of astacin proteases in the sea urchin genome that have been inferred from sequence predictions are listed in Fig. 2. All but two appear to be functional zinc metalloproteases because they contain all of the residues essential for catalytic activity: (1) the three histidine residues required to coordinate the catalytic zinc ion in the active center of the protease ($H_{92}E_{93}xxH_{96}xxGxxH_{102}$); (2) Y_{149} within the Met-turn, which also contains a conserved methionine (M147) (Stocker et al., 1993); and (3) the water-bound E_{93}, the general base in substrate hydrolysis (Yiallouros et al., 2002). The exceptions are SPU_005234, which lacks the conserved Met-turn (SxMxY) and SPU_017070, which lacks the required Y149 residue in the Met-turn (Fig. 2, in brackets). The 10 genes that are expressed during embryogenesis are marked by gene ID numbers shaded in lavender.

It is important to note that, with the exception of several genes that have been studied previously [*SpAN* (SPU_004113; Reynolds et al., 1992), *suBMP1* (SPU_007317, Hwang et al., 1994) and the *SpAN-like* genes, see below], the domain configurations listed in Fig. 2 are based solely on gene predictions

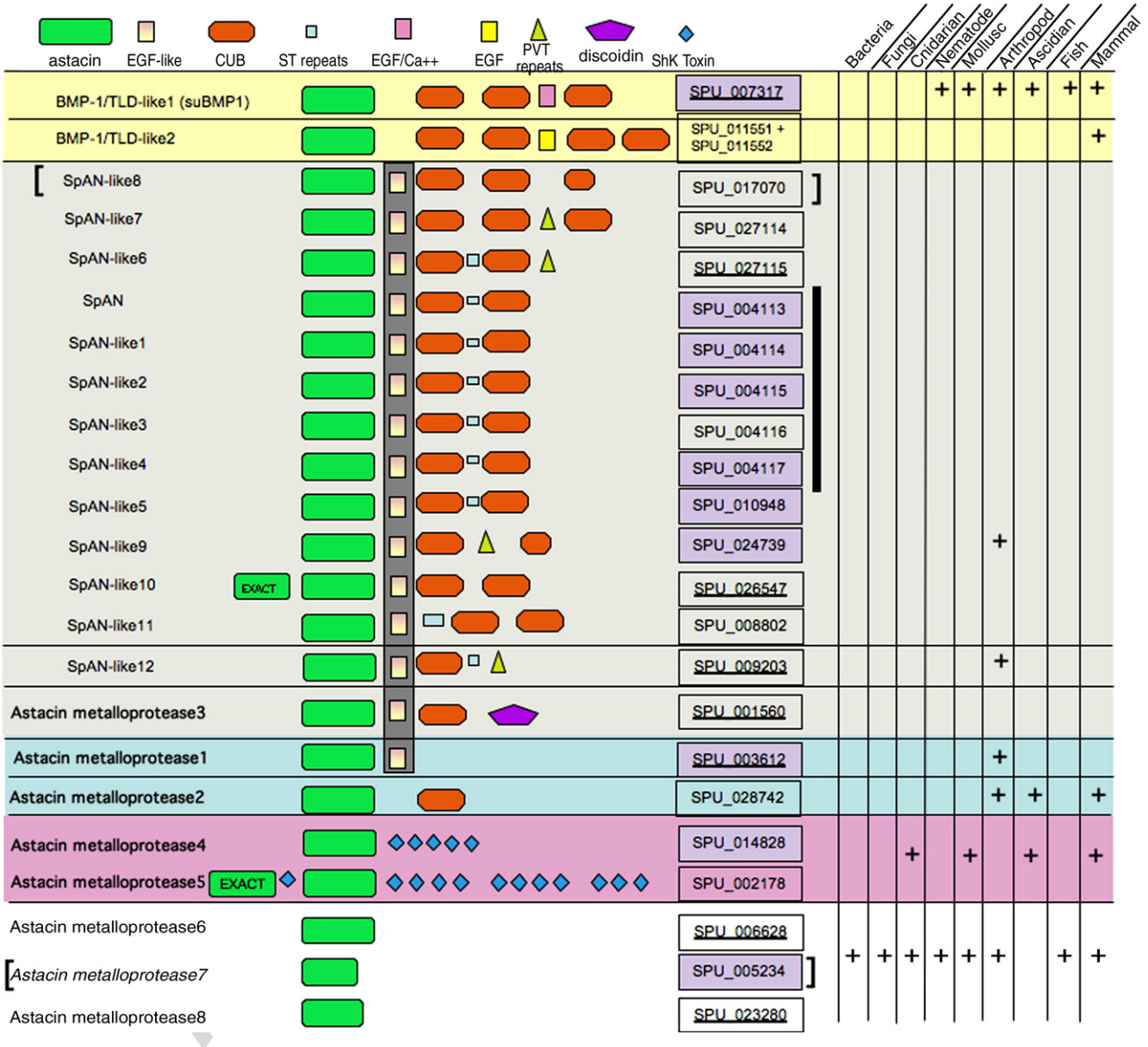


Fig. 2. The domain architectures of the 23 astacin gene predictions in the sea urchin genome are listed in order of decreasing complexity from top to bottom. The names at left reflect similarities to other astacin proteases based on domain architecture (symbols for different domain types are shown at the top). Gene ID numbers are listed to the right of the domain structures. These numbers refer to annotations of Glean3 predictions of the genome assembly as of 07/18/2005. Genes whose IDs are boxed in purple are expressed during embryonic development. The genes are grouped into 5 categories, indicated by yellow, gray, blue, pink and white backgrounds. Whether genes encoding sea urchin astacin proteases of different domain organization are present in the genomes of other organisms is indicated at right. The genes marked by brackets probably encode an inactive protease because required amino acid residues in the Met-turn are missing. Astacins included in the gray box share a cysteine-rich (EGF-like) domain immediately downstream from the protease domain and constitute the largest subgroup. Only nematodes have astacins containing a similar domain. Members of this group are named SpAN-like, after SpAN, the first one identified (Reynolds et al., 1992). Four SpAN-like predictions are clustered with SpAN, as indicated by the vertical black bar.

using the current genome assembly. Some of these may be incomplete because they reside near the ends of scaffolds (indicated by underlined *Gene IDs* in Fig. 2). In addition, some domains may have been included incorrectly. For example, in SPU_026547 and SPU_002178, the presence of two identical adjacent astacin protease domains is likely to result from an assembly error.

In another case (SPU_001560), the protein contains a discoidin domain downstream of a CUB domain. While this pair of domains can be found in many proteins, it has not been detected in astacin proteases. Whether the prediction is incorrect or whether this novel architecture is a sea urchin invention is not clear. Despite these uncertainties, sufficient conserved sequence has been identified to directly test the gene predictions by experimentation as well as to categorize the astacin genes into structural groups.

We have organized the sea urchin astacins into 5 structural classes, as indicated by shadings of different colors in Fig. 2. The white group contains representatives of the ancient, astacin protease domain-only proteases. Those in the pink group contain multiple ShKToxin domains-short regions bearing a conserved pattern of 6 cysteine residues characteristic of the primitive astacins—which are also found in cnidarians, nematodes and some vertebrates. The blue group representatives contain only one CUB (C) domain, while those in the yellow group have acquired multiple CUBs as well as EGF-like (E) repeats. This last group includes the two proteins most similar to BMP-1/TLD, although in vertebrates, flies and nematodes, the BMP-1/TLD proteins are characterized by additional CUB and EGF domains (CCECECC). These proteins are of major interest because their functions include processing extracellular matrix components, including collagen (reviewed by Sarras, 1996; Trackman, 2005), and regulating levels of the TGF-β family signaling ligands, BMP (Marques et al., 1997; Piccolo et al., 1997; Pappano et al., 2003) and GDF11 (Ge et al., 2005). Both BMP-1 and TLD activities are likely to be required during sea urchin embryogenesis as correct collagen assembly in the ECM is required for gastrulation (Wessel and McClay, 1987) and BMP signaling has been shown to be important in patterning cell fates along both the animal–vegetal and oral–aboral axes (Angerer et al., 2000). However, only one of the two sea urchin BMP-1/TLD-like proteases (SPU–007317; suBMP1; Hwang et al., 1994) is expressed in the embryo. Further support that this gene encodes a BMP-1/TLD protein that comes from phylogenetic analyses (see below; Fig. 4). These observations raise the question of whether suBMP1 fulfills both TLD and BMP-1 functions. There is precedent for this possibility: in the mouse, BMP-1 and TLD are encoded by the same transcription unit and are translated from alternatively spliced mRNAs (Takahara et al., 1994). Hwang et al. (1994) have reported that suBMP1 is expressed during the morphogenetic period between the blastula and pluteus stages. However, its relatively uniform spatial pattern of expression is not informative about its possible developmental roles and no direct tests of its function have been made. If this gene does not carry out both BMP-1 and TLD functions during embryogenesis, then the most likely candidates to execute them are those highlighted in gray.

An unusual feature of the *S. purpuratus* astacin protease repertoire is the expansion of a set having a distinctive structural organization, in which a cysteine-rich region (pink/yellow box) is immediately adjacent to the protease domain (Fig. 2, dark gray vertical shading). We will refer to this class as PEgfl (Protease, Egf-like). Although the cys-rich region resembles authentic EGF domains in the distribution of cysteine residues, it is identified as an EGF domain in only a subset of these predicted proteins using pFAM (http://www.sanger.ac.uk/Software/Pfam/) or CDART (Conserved Domain Architecture Retrieval Tool; NCBI). The only other PEgfl astacins are found in

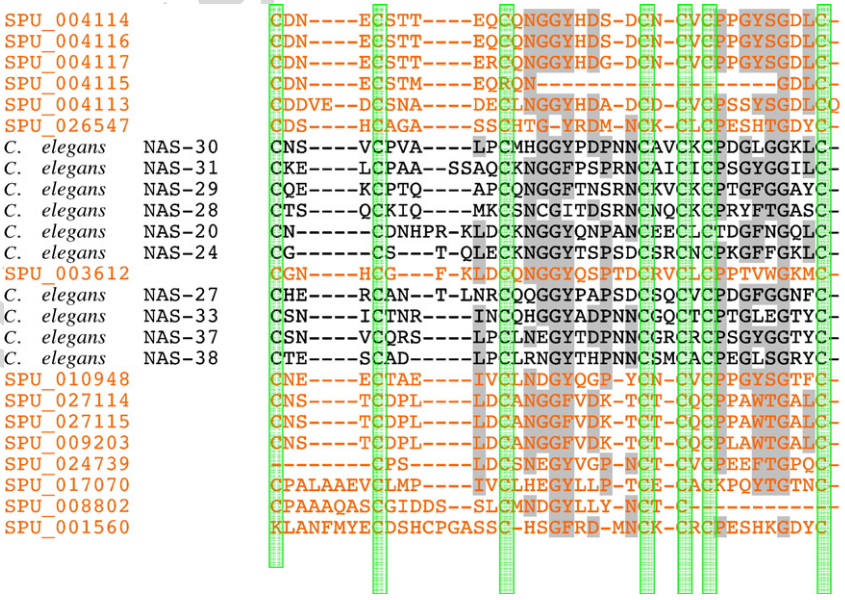


Fig. 3. ClustalW alignment of cys-rich, EGF-like domain sequences from *S. purpuratus* (orange letters) and nematodes (black letters). Areas are shaded when identical or very similar amino acids are present in the majority of cases. The 7 cysteines within this domain are boxed in green.

nematodes. That the Egf-like domain constitutes a functional domain is supported by the conservation in primary sequences within the cys-rich region in sea urchin and nematode astacins, which are shown aligned in Fig. 3. The fact that this unique structure is restricted to a basal deuterostome and a protostome is unusual. The most likely explanation is that coupling of the protease and EGF-like domains were separate events in echinoderms and nematodes, and the apparent similarity is due to convergent evolution. The possibility remains that this combination of domains appeared in the last common ancestor of protostomes and deuterostomes and was maintained only in these two groups. However, the apparent absence of this combination in other metazoans supports convergence.

We further analyzed the relationships among sea urchin astacins using neighbor-joining phylogenetic analysis of the sequences encoding the protease domains (Fig. 4). Most (9/11) of the P-Egfl astacins clustered in two groups, one of which consists of the SpAN-like proteins. These results show a correlation of protease domain similarity and the C-terminal domain architecture within the sea urchin astacins. Similar analyses using the protease domains of nematode PEgfl astacins (listed in Fig. 3) showed that they did not cluster with the corresponding sea urchin proteins (data not shown). This was also the case for SPU_017070 and SPU_003612. These observations also support the idea that the combination of an astacin protease domain and an EGF-like domain were separate events in echinoderms and nematodes. The duplications of these genes in both groups and their persistence as apparently functional genes suggest that this combination of domains was advantageous to these organisms. There appears to be a subgroup within the Egfl astacins that have acquired simple repeats of varying lengths that contain primarily proline, valine and threonine (PVT). The phylogenetic analysis also reveals that only one sea urchin protein (SPU_007317), discussed above, clusters with the BMP-1/TLD class and that one containing the ShKT domains (SPU_014828) is conserved with an ancient cnidarian protein (Heast4). In summary,

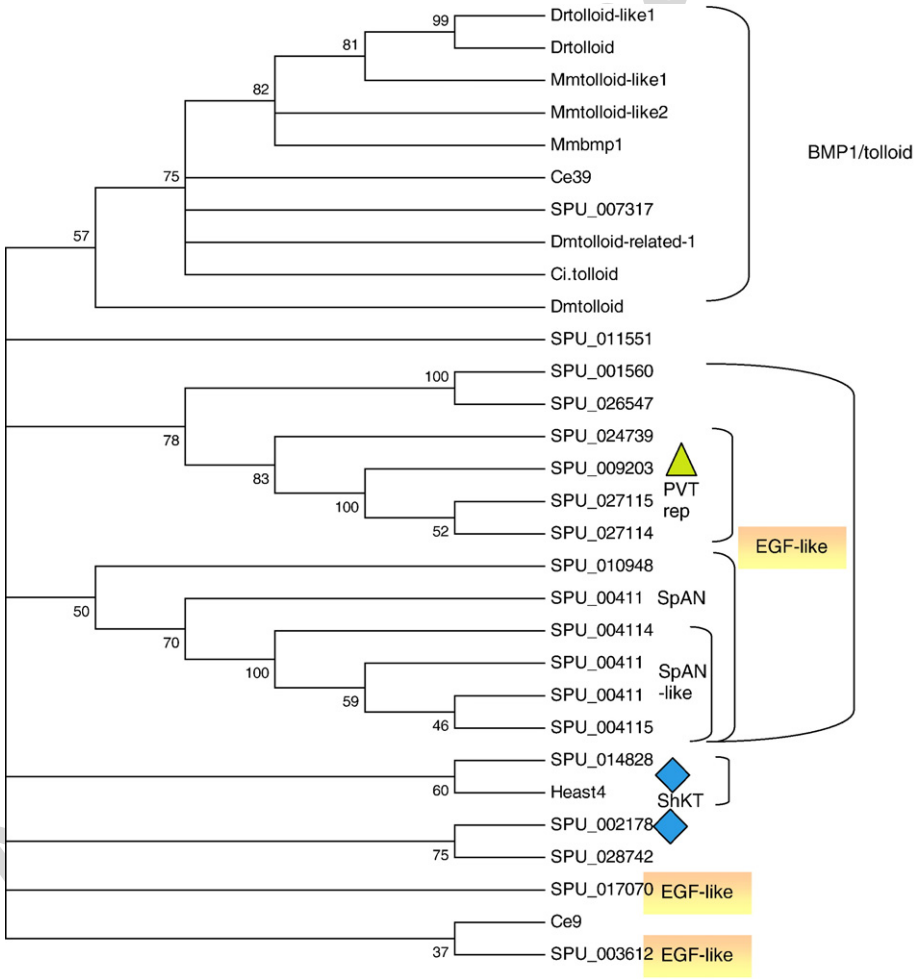


Fig. 4. Amino acid sequences of sea urchin astacin gene predictions were aligned with representative astacin protease domains using ClustalX (Thompson et al., 1997). The alignments were imported into MEGA 3.1 (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2001). The protease domains were aligned and neighbor-joining trees computed in MEGA, and represent 1000 bootstrap replicates. Colored symbols refer to domains described in Fig. 2 (green triangle, PVT repeats; blue diamond, ShKtoxin domains; pink/yellow rectangle, EGF-like domain immediately adjacent to the protease domain). Accession numbers for genes included in this analysis are: Drtolloid-like1 (NP_571085), Drtolloid (AAC60304), Mm tolloid-like1 (NP_033416), Mm tolloid-like2 (NP_036034), Mmbmp1 (NP_033885), CeNAS39 (NP_510672), Dmtolloid-related-1 (AAA70057), Citolloid (BAE06735), Dmtolloid (NP_524487), Heast4 (CAJ57450), Ce9 (NP_504293).

analysis of domain architectures, primary sequence similarities of the protease domain and the tiling microarray data indicate that the embryo expresses genes encoding a BMP/TLD-like molecule, an ancient ShKT domain-containing astacin protease and seven (at least) PEgfl astacins.

The prototypical member of the PEgsl astacins in *Strongylocentrotus purpuratus* is *SpAN*, a gene expressed in a highly dynamic and regulated pattern during early embryogenesis (Reynolds et al., 1992). *SpAN* transcripts accumulate transiently in most cells during blastula stages, but interestingly, not in those at the vegetal pole of the embryo that give rise to the mesoderm. SpAN protein is secreted to the apical extracellular matrix and preliminary studies suggest that its functions include modulating the structure of hyalin, an ECM component required for normal morphogenesis (E. Howard, L. Angerer, R. Angerer, unpublished observations). Unexpectedly, in the current assembly, *SpAN* (SPU_004113) is linked to 4 *SpAN-like* genes that are nearly identical to each other but only 70% identical to *SpAN* (SPU_004114, SPU_004115, SPU_004116, SPU_004117; marked by a vertical black bar in Fig. 2 and diagrammed in Fig. 5A). The sequence similarity is so high in both translated and untranslated regions that it is difficult to resolve whether these represent 4 different genes, as now indicated in the assembly, or whether several correspond to alleles that have been misassembled. Therefore, we will refer to these predictions as SpAN-like sequences rather than genes. At least two sequences (SPU_004114 and 004117) appear to be expressed during embryogenesis since well-matched ESTs in their 3′UTR regions exist (Fig. 5A, arrows). Semi-quantitative RT-PCR with primers against distinctly different sequences in the 3′UTRs show that SPU_ 004117, SPU_04116 and SPU_004114 mRNAs accumulate in a transient pattern during mesenchyme blastula stages (SPU_004117; Fig. 5B). Whether SPU_004116 sequences are expressed during embryogenesis is not clear. To determine where SpAN-like sequences are expressed in the sea urchin embryo, we carried out whole mount in situ hybridization using probes representing most of the coding regions since no sequence-specific probes could be designed that were of sufficient length to produce good in situ hybridization signals. The in situ signals at different stages confirm the RT-PCR results and reveal that SpAN-like sequences are expressed in an unusual pattern (Fig. 5C). Transcripts begin to accumulate in the early blastula in scattered cells that are primarily located in the animal hemisphere. As developmental proceeds, SpAN-like-positive cells become arranged in a more contiguous patch. The fact that these proteins have signal peptides and are likely to function extracellularly invites the speculation that they may be necessary for cells to move and coalesce near the animal pole. Thus far, SpAN-like function remains an open question because knockdowns using splice-blocking morpholinos targeting all the *SpAN-like* sequences have not been successful.

The genome sequence has revealed a surprising number of genes encoding astacin proteases in the sea urchin genome. The novel patterns of expression of several raise the possibility that they could have important developmental functions. Other genes worthy of study include additional *SpAN-like* genes (SPU_010948, 024739 and 03612) and a conserved and ancient ShKT-containing astacin (SPU_014828) that are expressed in the embryo (Samanta et al., in press).

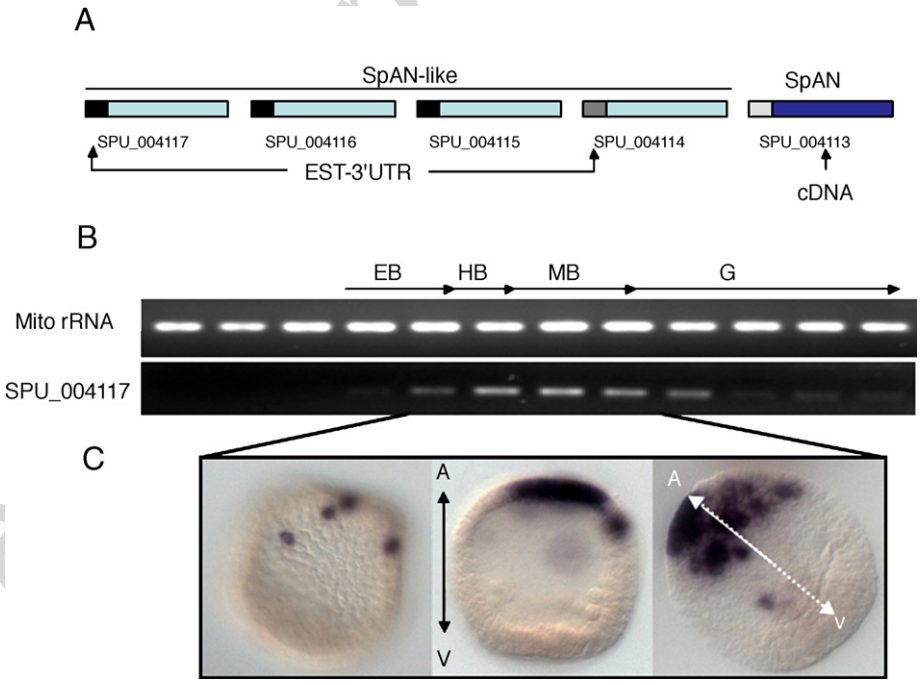


Fig. 5. (A) Schematic of the *SpAN/SpAN-like* cluster. The 5′−3′ direction of the coding strand is from right to left. Full-length and partial cDNA sequences have been recovered for *SpAN* (Reynolds et al., 1992) and *SpAN-like* genes, respectively. (B) RT-PCR products representing SPU_004117 mRNAs present during egg to gastrula stages. Estimates of total RNA levels at different stages are provided by measurements of mitochondrial 12s RNA. EB (early blastula); HB (hatching blastula); MB (mesenchyme blastula); G (gastrula). (C) Whole mount in situ hybridization was carried out as described by Minokawa et al. (2004) using a cDNA probe that represents mRNAs from any of the *SpAN-like* genes.

Matrix metalloproteinases

The matrix metalloproteinases constitute one of the major families of proteinases that act on the extracellular matrix. There are 24 human MMPs, and homologues have been found in most animals as well as plants and algae. The minimal MMP consists of a prodomain followed by a catalytic domain (Vu and Werb, 2000). The prototype MMP is a minimal MMP linked to a set of four hemopexin repeats. Examination of the available animal genomes indicates that each contains a gene with this architecture, including those with a single MMP gene, such as the cnidarians and molluscs (Fig. 6A). Among the vertebrate MMPs, the number of hemopexin repeats varies and some also have transmembrane domains (MT-MMPs) or fibronectin-like type II repeats.

We found that there are at least 26 genes encoding MMPs in the sea urchin genome (Fig. 6A). All but two of the predicted peptidase M10 domains in the MMP gene models contain the HEXXH zinc-binding site amino acid sequence. In SPU_023505, the critical catalytic residue, glutamate, has been replaced with glutamine and is likely to be non-functional. The protease domain of SPU_005376 is truncated and probably an incomplete gene model. Thirteen genes include the N-terminal PG binding domain (shaded yellow in Fig. 6A). Six of these encode proteins with the prototypical domain architecture. In addition, there are genes encoding proteins with zero, one, two

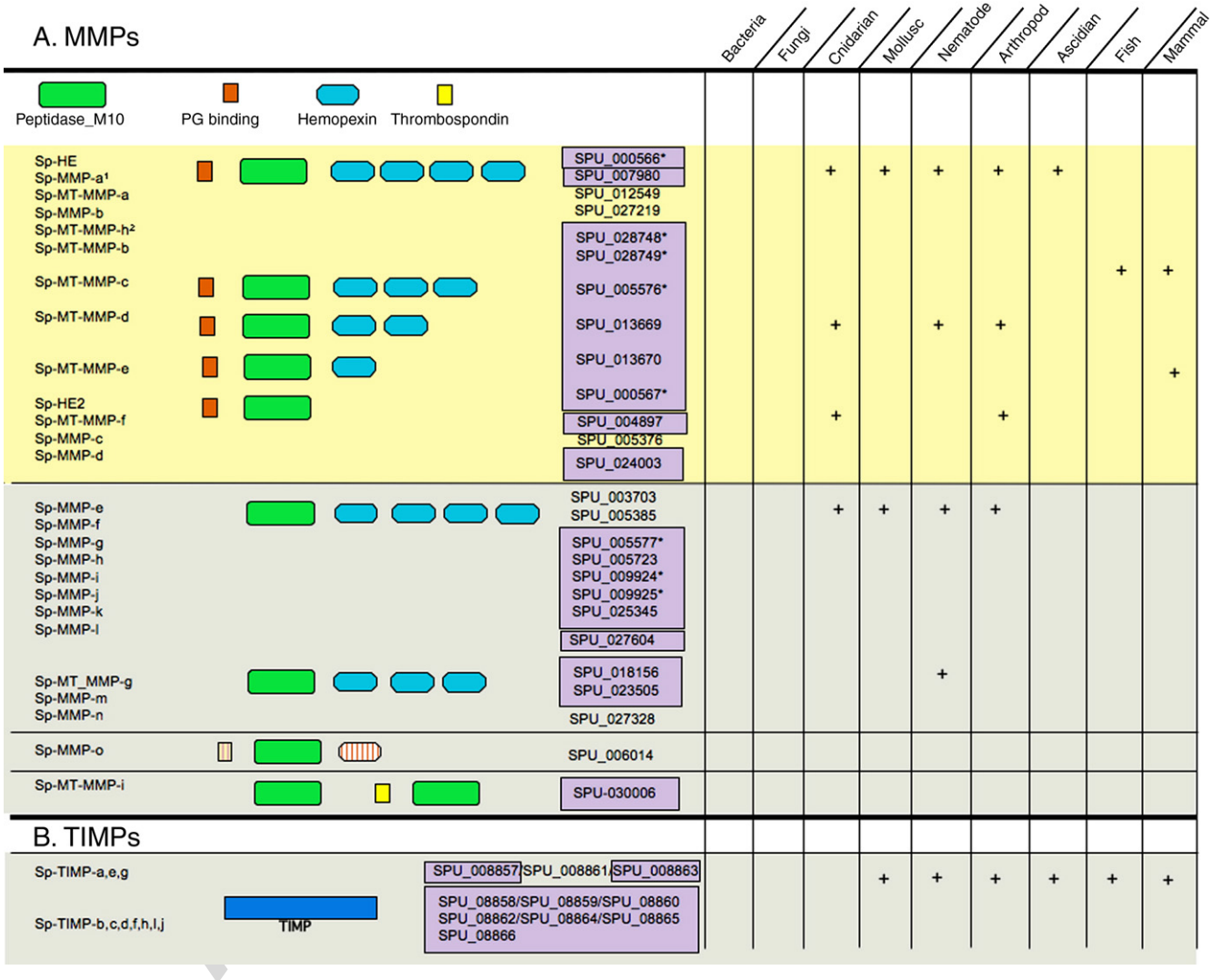


Fig. 6. The domain architecture of the 26 MMP (A) and 10 TIMP (B) gene predictions in the sea urchin genome (genome assembly 7/18/2005) are listed. (A) The first group of MMP genes contains those with an N-terminal PG domain, followed by a peptidase domain and varying numbers of hemopexin repeats. The second group lacks the N-terminal domain. Two unusual MMP gene predictions are shown at the bottom of the list. A key to the domains is shown at the top. The names and gene ID numbers of these genes are given in columns on the left and right, respectively. Genes in lavender boxes are expressed during embryonic development. The columns on the far right list other phylogenetic groups that were examined for genes with a similar domain architecture. All domain architectures found in the sea urchin genome are shown; however, other groups may have genes in this family with different domain architectures. If a gene with the same domain architecture as sea urchins was found in a given phylogenetic group, a plus was placed in the column. An asterisk next to a gene ID number indicates that the gene is found in a cluster. Dashed boxes indicate a domain with a low prediction score. (B) TIMP genes are listed as above (indicated with a blue background). All of the TIMP genes are found in one cluster.

or three hemopexin repeats following the catalytic domain. Similar domain architectures are found in other vertebrate genomes and in the nematode, but is rare in other animals (Fig. 6A). In some cases, the reduced number of hemopexin domains may result from incomplete gene predictions in the current genome assembly, but for others, this is unlikely.

The sea urchin genome also contains MMPs that lack the N-terminal, proteoglycan (PG)-like binding domain (shaded gray in Fig. 6A). These have either three or four hemopexin repeats. The latter class is found in nematodes, arthropods and vertebrates, but not in ascidians (Fig. 6A), while the former have only been observed in nematodes. There are also two unusual MMP-like genes in the sea urchin genome. One has a normal catalytic domain, but the N- and C-terminal domains are only weakly similar to the respective PG binding-like and hemopexin domains in other MMPs (Fig. 6A; striped domains). The second has two catalytic domains flanking a PG binding-like domain and is likely due to an assembly error.

We used BLAST to find the most likely homologues for each of the sea urchin genes. All gave significant alignment scores with vertebrate genes in the MMP family. However, when amino acid sequence alignments using ClustalW and MEGA were used to construct phylogenetic trees in MEGA using the neighbor-joining method, the sea urchin genes clustered together, separate from vertebrate MMP groups (Fig. 7). This analysis suggests that MMP genes found in the last common ancestor to vertebrates and echinoderms underwent substantial duplication and divergence following separation of the two groups. The phylogenetic evaluation indicates that we cannot assign homology to individual MMPs, and therefore we have named these genes Sp-MMP.a-n.

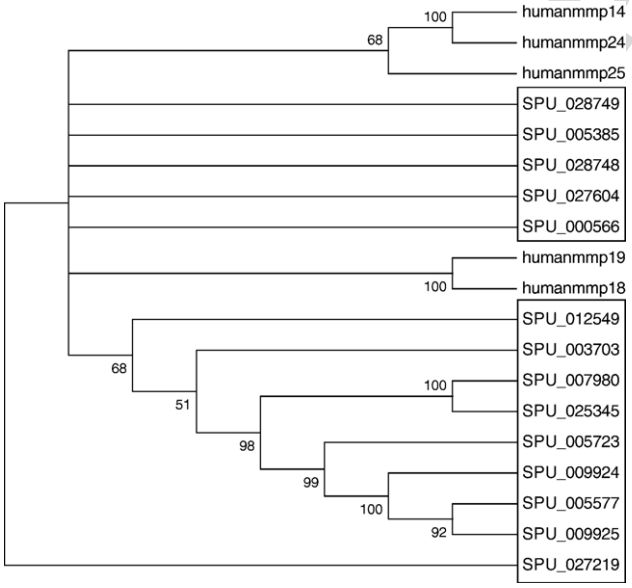


Fig. 7. Amino acid sequences of sea urchin MMP gene predictions were aligned with representative human MMP and MT-MMP genes using ClustalX (Thompson et al., 1997). They were imported into MEGA (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2001) and domains were individually aligned using Clustal. Neighbor-joining trees were computed using MEGA and represent 1000 bootstrap replicates. Accession numbers for genes used in this analysis: Human 14 (AAV40837), Human 18 (CAA69913), Human 19 (NP067387), Human 24 (AAH47614), Human 25 (NP071913).

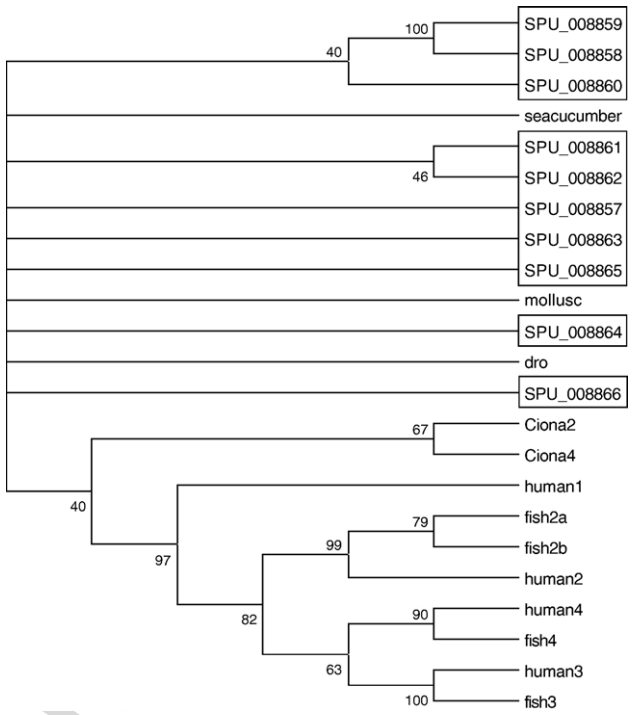


Fig. 8. Amino acid sequences of the sea urchin TIMP gene predictions were aligned using ClustalX (Thompson et al., 1997) with representative TIMP genes from both deuterostomes and protostomes whose genomes have been sequenced. They were imported into MEGA (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2001) and neighbor-joining trees were computed using MEGA and represent 1000 bootstrap replicates. Accession numbers for genes included in this analysis are: Human 1 (NP003245), Human 2 (NP003246), Human 3 (AAH14277), Human 4 (NP003247), Fugu 2a (BAE026264), Fugu 2b (BAE02620), Fugu 3 (BAE02621), Fugu 4 (AAO17737), Ciona 2 (Ci0100139932), Ciona 4 (Ci0100139906), Drosophila (NP731461.1), Sea cucumber (AAK61535.1).

One of the MMP gene types that may have been selectively amplified in echinoderms encodes membrane-tethered (MT)-MMPs. Homology searches using BLAST placed some sea urchin MMP genes as most similar to vertebrate MT-MMP genes. Using the TMpred prediction algorithm (Hofmann and Stoffel, 1993), we confirmed that at least 9 predictions have putative transmembrane domains. If they are MT-MMPs, this is nearly twice the number found in vertebrates. The genes that may encode membrane-bound MMPs are named Sp-MT-MMP. a-n (Fig. 6A).

There are at least five pairs of MMP genes adjacent to one another on genomic scaffolds (indicated with asterisks on Fig. 6A). Genes in these clusters do not group together in phylogenetic analyses (Fig. 7) and do not all share the same domain architecture. Consequently, they are not haplotypes placed adjacent to one another due to assembly errors. It is possible that that there are larger clusters of MMP genes, as is the case in mammalian genomes, since some gene pairs are near the ends of their scaffolds in the current assembly.

Of the 26 MMP genes identified, 19 are expressed in the embryo based on microarray data (Samanta et al., in press).

These are indicated by lavender boxes around the Gene IDs in Fig. 6. Based on EST data derived from cDNA libraries, they are expressed as early as 7 h and as late as larval stages. Eleven MMP genes are expressed at blastula (42%), and eight are expressed in primary mesenchyme cells (31%).

The activities of some MMPs are regulated by inhibitory proteins (tissue inhibitors of metalloproteases; TIMPs). Mammals have four TIMP genes, each on a separate chromosome. Each has a specific pattern of expression and differs in its ability to inhibit different MMPs (Brew et al., 2000; Bode and Maskos, 2003). Surprisingly, the sea urchin genome contains 10 TIMP genes (Fig. 6B, shaded in blue), and all of them are arranged in the same orientation in a cluster on a single genomic scaffold. Varying numbers of TIMP genes are found in molluscs (1), nematodes (1), *Drosophila* (1), *Ciona* (2), zebrafish (4) and mammals (4) (Fig. 6B). Nine of the 10 TIMP genes are expressed in the embryo based on microarray data (Samanta et al., in press). ESTs for six of the ten have been found in blastula cDNAs and four sequences are expressed in primary mesenchyme cells.

Alignment of the TIMP proteins in these groups of animals with those in the sea urchin genome using ClustalW followed by neighbor-joining analysis using MEGA resulted in the relationships shown in Fig. 8. All of the chordate TIMP genes group together, but the sea urchin TIMP genes fall outside that group. The phylogenetic analysis suggests the following model for TIMP gene evolution. It appears that the protostome–deuterostome ancestor had a single TIMP gene. In the chordate line, this gene duplicated once in *Ciona* following its divergence from vertebrates. In humans, four TIMP genes arose by gene duplications. Fish have two TIMP2-like genes, likely due to an independent duplication, one TIMP3 and TIMP4, but not TIMP1. In sea urchins, tandem duplications have given rise to 10 adjacent genes. The most recent duplications generated the cluster of SPU_008859 through SPU_008862. Deeper relationships are unresolved by this analysis.

Due to the independent duplications in echinoderms and vertebrates of the MMP and TIMP genes, it is not possible to determine orthology. It is interesting to note that expansion and divergence of some MMP genes and the TIMP genes occurred in the line leading to sea urchins. These two sets of genes encode interacting proteins raising the possibility that co-evolution played a role in this process. It is also interesting that a large proportion of both the MMP and TIMP genes are expressed in the migratory primary mesenchyme cells that give rise to the larval skeleton.

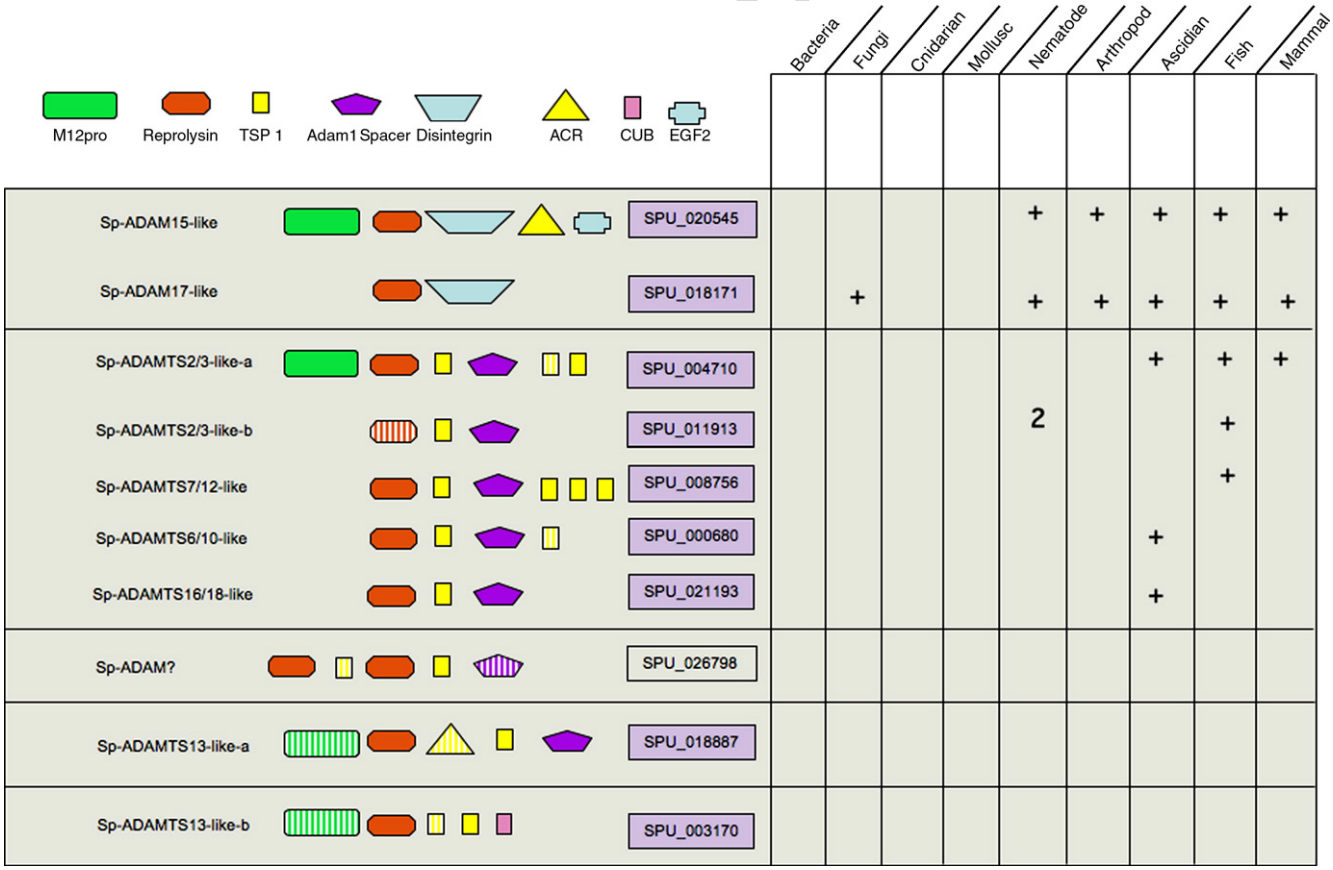


Fig. 9. The domain architecture of the ADAM gene predictions (indicated by a yellow background) and ADAMTS gene predictions (indicated by a gray background) in the sea urchin genome (genome assembly 7/18/2005) are listed. A key to the domains is shown at the top. The names and gene ID numbers of these genes are given in columns on the left and right, respectively. Genes in lavender boxes are expressed during embryonic development. The columns on the right list other phylogenetic groups that were examined for genes with similar domain architecture. All domain architectures found in the sea urchin genome are shown; however, other groups may have genes in this family with different domain architectures. If a gene with the same domain architecture as sea urchins was found in a given phylogenetic group, a plus was placed in the column. Dashed boxes indicate a domain with a low prediction score.

ADAM and ADAMTS genes

ADAM and ADAMTS genes are highly conserved in chordates. The sea urchin genome also contains genes similar to some, but not all, members in both of these groups. This is particularly true for the ADAMs since only two genes have been detected. Sea urchins also have fewer ADAMTS genes than other deuterostomes, but most of the chordate groups are represented. All the ADAM and ADAMTS genes contain sequences encoding the consensus HEXXH zinc-binding site and are likely to be proteolytically active.

Four sea urchin gene predictions corresponding to two pairs of alleles contain complete metalloprotease and disintegrin domains and therefore were considered to encode ADAMs (Fig. 9, shaded yellow). Based on reciprocal BLAST searches, they were named Sp-ADAM15-like (SPU_020545) and Sp-ADAM17-like (SPU_018171). Sp-ADAM15-like contains reprolysin family propeptide, reprolysin, disintegrin, ADAM-CysRich (ACR) and EGF2 domains in addition to the metalloprotease domain, and was originally described by Rise and Burke (2002). The Sp-ADAM17-like sequences contain only Reprolysin and Disintegrin domains. Genes with similar domain structures can be found in all chordates as well as some protostomes and fungi (Fig. 9).

ADAM17 was the first ADAM implicated in ectodomain shedding, and is essential to mammalian development (Huovila et al., 2005). Further studies indicate that it performs many important functions in adult mammals including the inflammatory response by regulating growth factors and hormone signaling (Huovila et al., 2005; Handsley and Edwards, 2005). It is of little surprise, therefore, that this gene appears to be conserved in highly diverse taxa. Sequences similar to *ADAM17* have been found in vertebrates, *Ciona*, sea urchin, *Drosophila* and nematodes (Fig. 9).

Phylogenetic analysis carried out as described for MMPs and TIMPs generated a neighbor-joining tree showing that the two sea urchin ADAM sequences fall within two clades. Sp-ADAM17-like (SPU_018171) clearly groups with other ADAM17 genes from both protostomes and deuterostomes (Fig. 10). The position of Sp-ADAM15-like (SPU_020545) is less clear. It falls within the clade of ADAM12/15/33, and seems most closely related to a *Ciona* sequence (*ADAM12/15/19/33*) (Fig. 10). However, these relationships are not strongly supported. The vertebrate genes in this clade have undergone considerable duplication and divergence following divergence of vertebrates from ascidians and echinoderms, which makes it difficult to determine the relationships of the sea urchin and Ciona sequences to the vertebrate genes.

ADAM10 is highly conserved in nematodes, *Ciona* and vertebrates where it has been shown to have critical functions. For example, this protease is required for Notch signaling and *ADAM10*-knockout mouse embryos are not viable (Huovila et al., 2005); it has been also implicated in neuronal development in adult mice (Huovila et al., 2005).

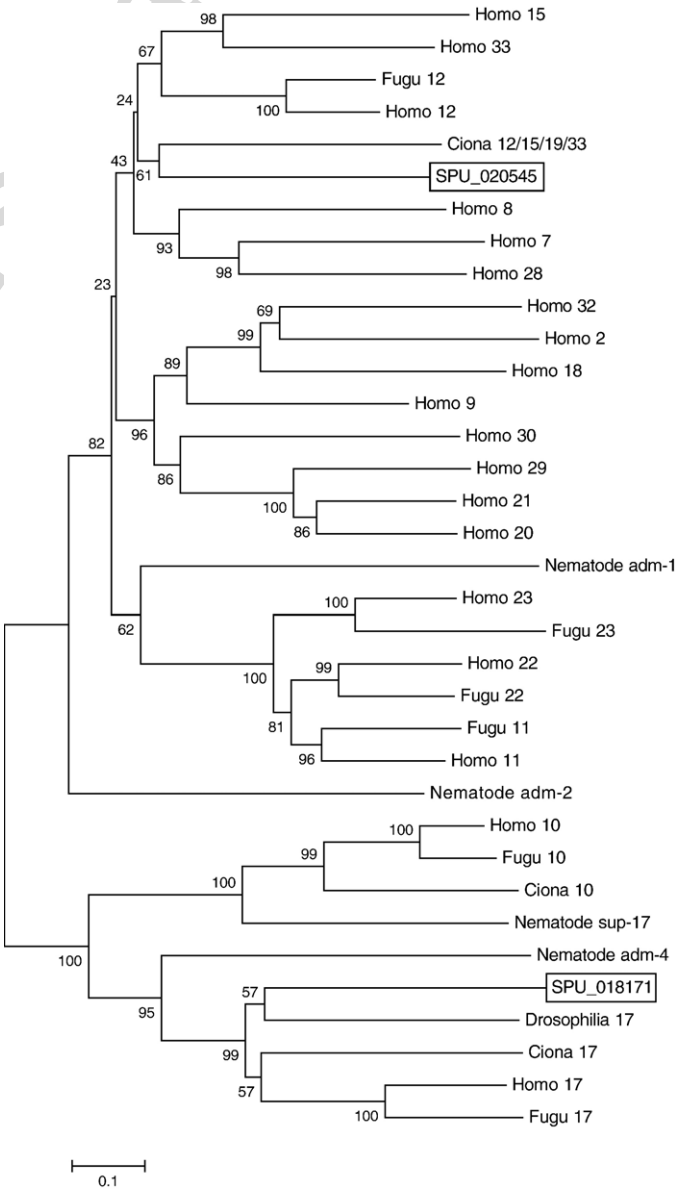
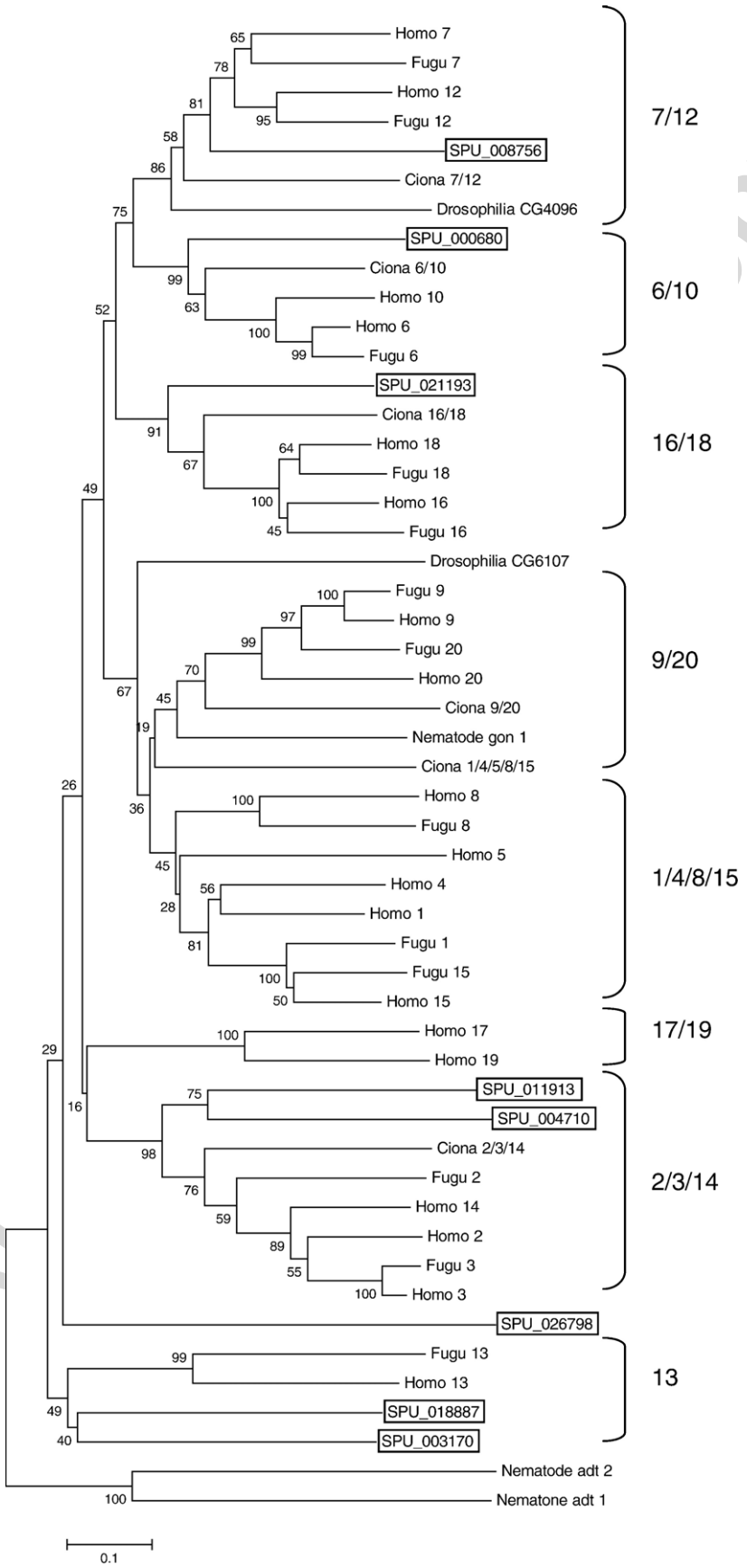


Fig. 10. Amino acid sequences of the sea urchin ADAM gene predictions were aligned using ClustalX (Thompson et al., 1997) with representative ADAM genes from both deuterostomes and protostomes whose genomes have been sequenced. They were imported into MEGA (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2001) and domains were individually aligned using ClustalW. Neighbor-joining trees were computed using MEGA and represent 1000 bootstrap replicates. Sequences for *Homo sapiens* and *Drosophilia melanogaster* were collected from Genbank: Homo 2, CAB40813; Homo 7, NP_003808; Homo 8, NP_001100; Homo 9, NP_003807; Homo 10, NP_001101; Homo 11, NP_002381; Homo 12, AAQ89237; Homo 15, AAS73000; Homo 17, NP_003174; Homo 18, AAQ88687; Homo 20, AF029899; Homo 21, AAI09025; Homo 22, AF073291; Homo 23, CAC20565; Homo 28, AAD25099; Homo 29, AF171929; Homo 30, AF171932; Homo 32, BC030014; Homo 33, AB055891; Drosophilia 17, AAO53296. Nematode sequences were collected from WormBase (www.Wormbase.org): (Nematode sup-17, DY3.7; Nematode adm-4, ZK154.7; Nematode adm-2, C04A11.4; Nematode adm-1, Y37D8A.13). Sequences for *Fugu* were collected from the Joint Genome Institute (http://genome.jgi-psf.org/cgi-bin/runAlignment?db=Takru4_and_program=tblastn_and_dataLib=fugu_and_email=_and_advanced=1 using version 4.0): Fugu 10, e_gw2.14.131.1; Fugu 11, e_gw2.29.32.1; Fugu 12, e_gw2.76.83.1; Fugu 17, e_gw2.268.6.1; Fugu 22, e_gw2.48.128.1; Fugu 23, e_gw2.100.126.1. Sequences for *Ciona* were also collected from Joint Genome Institute (http://genome.jgi-psf.org/cgi-bin/runAlignment?db=ciona4_and_advanced=1): Ciona 10, ci0100139599; Ciona 12/15/33, ci0100140827; Ciona 17, ci0100152455.


Homo 7
Fugu 7
Homo 12
Fugu 12
SPU_008756
Ciona 7/12
Drosophilia CG4096
7/12
SPU_000680
Ciona 6/10
Homo 10
Homo 6
Fugu 6
6/10
SPU_021193
Ciona 16/18
Homo 18
Fugu 18
Homo 16
Fugu 16
16/18
Drosophilia CG6107
Fugu 9
Homo 9
Fugu 20
Homo 20
Ciona 9/20
Nematode gon 1
Ciona 1/4/5/8/15
9/20
Homo 8
Fugu 8
Homo 5
Homo 4
Homo 1
Fugu 1
Fugu 15
Homo 15
1/4/8/15
Homo 17
Homo 19
17/19
SPU_011913
SPU_004710
Ciona 2/3/14
Fugu 2
Homo 14
Homo 2
Fugu 3
Homo 3
2/3/14
SPU_026798
Fugu 13
Homo 13
SPU_018887
SPU_003170
13
Nematode adt 2
Nematone adt 1
0.1

Surprisingly, it appears to have been lost in sea urchins. Overall, the number and diversity of sea urchin ADAM genes is much reduced compared to vertebrates, but similar to *Ciona* and protostomes.

Rise and Burke (2002) described that SpADAM (SpADAM15-like; SPU_020545) is expressed during cleavage in all cells, then later by vegetal plate cells, mesenchyme cells muscles and neurons. SPU_018171 is expressed based on microarray analysis (Samanta et al., in press) and is found in the primary mesenchyme EST database.

Eight sea urchin gene predictions contain the collection of domains characteristic of ADAMTS proteins [disintegrin, metalloproteinase and thrombospondin (TS)] (Fig. 9, shaded gray). In addition, all except one contain an ADAM1 spacer. Four sequences differ only in the number of TS motifs and one contains an additional propeptide domain. The architecture of these five sequences is found in other deuterostomes, but not outside that group. Surprisingly, three of these sequences contain additional domains not yet found in other ADAMTS sequences. These include an additional reprolysin domain in SPU_026798, and a CUB domain in SPU_003170 and an ACR domain in SPU_018887 usually found in ADAMs rather than ADAMTS gene sequences (Fig. 9).

All but one of the sea urchin ADAMTS genes, except for SPU_26798, are expressed in the embryo based on microarray analysis (Samanta et al., in press). Based on searches of EST databases, SPU_008756 is represented in blastulae and larvae, SPU_004710 is maternal and in primary mesenchyme and blastula-stage embryos and SPU_003170 is found in blastulae, primary mesenchyme and larvae.

The eight predicted sea urchin ADAMTS gene sequences were also aligned with representative genes from other taxa using Clustal and MEGA and used to construct a neighbor-joining tree using MEGA, which reveals several trends in the evolution of ADAMTS sequences. There are seven distinct groups of ADAMTS genes that contain deuterostome sequences (Fig. 11). Most of these are deuterostome-specific, supporting the hypothesis that gene duplications have occurred in deuterostomes following their divergence from protostomes (Nicholson et al., 2005). Our analysis sheds some light on the timing of those duplications. We refer to clades based on the vertebrate genes found in each group. It is clear that clades 6/10 and 7/12 are closely related (Nicholson et al., 2005), but since there are sea urchin genes in each of these clades, it appears that the duplication event leading to these two groups occurred before sea urchins diverged from other deuterostomes. There is also a single sea urchin gene that clusters with the ADAMTS 16/18 group, indicating the ancestor to all deuterostomes had a copy of this gene. The absence of protostome genes in these groups indicates that they arose following the divergence of protostomes and deuterostomes. There are two sea urchin genes that group with the ADAMTS2/3/14 clade. However, the sea urchin genes fall outside the other deuterostome group, suggesting that there has been an independent duplication of the ancestral gene in the echinoderm lineage. Clearly, these four clades of ADAMTS genes diverged prior to the separation of chordates and echinoderms. There are no sea urchin sequences in the 9/20 or the 1/4/5/8/15 clades, and there are protostome genes associated with these combined groups. Previous studies have shown that the ADAMTS1/4/8/15 clade is a chordate-specific group of genes (Nicholson et al., 2005). Sea urchins have apparently lost the gene that was the progenitor to these two clades. Finally, ADAMTS13 is unique among ADAMTS proteins in that it contains an essential CUB domain. Thus far it has only been found in vertebrates, but there may be sea urchin ADAMTS13 homologues because two sea urchin sequences cluster with the *Fugu* and human ADAMTS13 sequences and one of them (SPU_000317) also has a CUB domain (Figs. 9 and 11). With the exception of SPU_026798, each of the ADAMTS genes in the sea urchin genome has clear homology with a specific group of chordate genes. This suggests that a deuterostome ancestor had a minimum of five ADAMTS genes.

Concluding remarks

Metalloproteases are required for normal development of the sea urchin embryo since inhibition of their activity disrupts morphogenesis. The idea that they have interesting developmental roles is supported by expression data for six previously studied metalloprotease genes in the BMP-1/TLD, MMP and ADAMs families that are highly regulated in time and space during early sea urchin development. These observations predict that additional members of this class of proteases will have important developmental functions. Now, through examination of the Glean3 set of predictions in the 7.18.2005 sea urchin genome assembly, we have identified 56 (or nearly 10 times as many) of these kinds of metalloprotease genes. Sixty

Fig. 11. Amino acid sequences of the sea urchin ADAMTS gene predictions were aligned using ClustalX (Thompson et al., 1997) with representative ADAM genes from both deuterostomes and protostomes whose genomes have been sequenced. They were imported into MEGA (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2001) and domains were individually aligned using ClustalW. Neighbor-joining trees were computed using MEGA and represent 1000 bootstrap replicates. Brackets on the right indicate groups of related genes. The bracket numbers indicate the vertebrate genes found in each group. Sequences for *Homo sapiens* and *Drosophilia melanogaster* were collected from Genbank: Homo 1, NP–008919; Homo 2, O95450; Homo 3, O15072; Homo 4, O75173; Homo 5, Q9UNA0; Homo 6, NP–922932; Homo 7, AAH61631; Homo 8, AAH89435; Homo 9, NP–891550; Homo 10, NP_112219; Homo 12, P58397; Homo 13, CAI17258; Homo 14, NP_542453; Homo 15, NP–620686; Homo 16, Q8TE57; Homo 17, NP–620688; Homo18, NP–955387; Homo 19, CAC84565; Homo 20, AAO15766; Drosophilia, CG4096; Drosophilia CG6107. Nematode sequences were collected from WormBase (www.Wormbase.org): Nematode gon-1, F25H8.3; Nematode adt-2, F08C6.1a.1; Nematode adt-1, C02B4.1. Sequences for *Fugu* were collected from Joint Genome Institute (http://genome.jgi-psf.org/): Fugu 1, e–gw2.6.372.1; Fugu 2, e–gw2.51.136.1; Fugu3, e–gw2.4.471.1; Fugu 6, e–gw2.27.235.1; Fugu 7, e–gw2.1.134.1; Fugu 8, e_gw2.144.133.1; Fugu 9, e–gw2.33.213.1; Fugu 12, e–gw2.50.400.1; Fugu 13, e–gw2.18.227.1; Fugu 15, e–gw2.144.132.1; Fugu 16, e–gw2.22.32.1; Fugu 18, e–gw2.92.52.1; Fugu 20, e–gw2.2.66.1. Sequences for *Ciona* were also collected from Joint Genome Institute (http://genome.jgi-psf.org/): Ciona 2/3/14, ci0100146117; Ciona 6/10, ci0100146470; Ciona 7/12, ci0100132574; Ciona 9/20, ci0100132719; Ciona 1/4/5/8/15, ci0100137065; Ciona 16/18, ci0100138085.

percent of them are expressed during sea urchin embryogenesis. A new subfamily of metalloprotease genes in the BMP-1-TLL group has emerged, and one of these has a novel expression pattern in the animal pole domain of sea urchin blastulae. Surprisingly, a large fraction of the MMPs appear to be membrane-tethered forms and the number of TIMP genes has more than doubled compared to vertebrates through gene duplications. Lastly, while the sea urchin has only two ADAMS genes, genes encoding a structurally diverse set of ADAMTS genes have been uncovered. The results indicate that expansion of the ADAMS genes occurred only in vertebrates, while the initial expansion of the ADAMTS genes occurred in the ancestor to deuterostomes. Analysis of the genome sequence has opened up a new, huge area of investigation into the regulation of the extracellular environment during sea urchin embryo development.

Acknowledgments

We thank Erica Sodergren, George Weinstock, Lan Zhang and other members of the sea urchin sequencing project at Baylor for providing the sequence and annotation analysis tools, and the tremendous efforts of Eric Davidson and members of his laboratory that made the sea urchin genome sequence a reality. We would also like to thank Bob Angerer for useful comments and Jim Garey for helpful discussions on phylogenetics and evolution. This work was supported by funds from the intramural program of the National Institute for Dental and Craniofacial Research in the National Institutes of Health.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2006.07.046.

References

Adelson, D.L., Humphreys, T., 1988. Sea urchin morphogenesis and cell-hyalin adhesion are perfurbed by a monoclonal antibody specific for hyalin. Development 104, 391–402.

Angerer, L.M., Oleksyn, D.W., Logan, C.Y., McClay, D.R., Dale, L., Angerer, R.C., 2000. A BMP pathway regulates cell fate allocation along the sea urchin animal–vegetal embryonic axis. Development 127, 1105–1114.

Blobel, C.P., 2005. ADAMs: key components in EGFR signalling and development. Nat. Rev. Mol. Cell Biol. 6, 32–43.

Bode, W., Maskos, K., 2003. Structural basis of the matrix metalloproteinases and their physiological inhibitors, the tissue inhibitors of metalloproteinases. Biol. Chem. 384, 863–872.

Brew, K., Dinakarpandian, D., Nagase, H., 2000. Tissue inhibitors of metalloproteinases: evolution, structure and function. Biochim. Biophys. Acta 1477, 267–283.

Burke, R.D., Myers, R.L., Sexton, T.L., Jackson, C., 1991. Cell movements during the initial phase of gastrulation in the sea urchin embryo. Dev. Biol. 146, 542–557.

Duboc, V., Rottinger, E., Besnardeau, L., Lepage, T., 2004. Nodal and BMP2/4 signaling organizes the oral–aboral axis of the sea urchin embryo. Dev. Cell 6, 397–410.

Duboc, V., Rottinger, E., Lapraz, F., Besnardeau, L., Lepage, T., 2005. Left–right asymmetry in the sea urchin embryo is regulated by nodal signaling on the right side. Dev. Cell 9, 147–158.

Emily-Fenouil, F., Ghiglione, C., Lhomond, G., Lepage, T., Gache, C., 1998. GSK3beta/shaggy mediates patterning along the animal–vegetal axis of the sea urchin embryo. Development 125, 2489–2498.

Ge, G., Hopkins, D.R., Ho, W.B., Greenspan, D.S., 2005. GDF11 forms a bone morphogenetic protein 1-activated latent complex that can modulate nerve growth factor-induced differentiation of PC12 cells. Mol. Cell. Biol. 25, 5846–5858.

Ghiglione, C., Lhomond, G., Lepage, T., Gache, C., 1994. Structure of the sea urchin hatching enzyme gene. Eur. J. Biochem. 219, 845–854.

Handsley, M.M., Edwards, D.R., 2005. Metalloproteinases and their inhibitors in tumor angiogenesis. Int. J. Cancer 115, 849–860.

Hofmann, K., Stoffel, W., 1993. TMbase-A database of membrane spanning protein segments. Biol. Chem. Hoppe-Seyler 374, 166.

Huggins, L., Lennarz, W., 2001. Inhibitors of procollagen C-terminal proteinase block gastrulation and spicule elongation in the sea urchin embryo. Dev. Growth Differ. 43, 415–424.

Huovila, A.P., Turner, A.J., Pelto-Huikko, M., Karkkainen, I., Ortiz, R.M., 2005. Shedding light on ADAM metalloproteinases. Trends Biochem. Sci. 30, 413–422.

Hwang, S.P., Partin, J.S., Lennarz, W.J., 1994. Characterization of a homolog of human bone morphogenetic protein 1 in the embryo of the sea urchin, *Strongylocentrotus purpuratus*. Development 120, 559–568.

Ingersoll, E.P., Wilt, F.H., 1998. Matrix metalloproteinase inhibitors disrupt spicule formation by primary mesenchyme cells in the sea urchin embryo. Dev. Biol. 196, 95–106.

Ingersoll, E.P., Pendharkar, N.C., 2005. Characterization and expression of two matrix metalloproteinase genes during sea urchin development. Gene Expr. Patterns 5, 727–732.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244–1245.

Lepage, T., Ghiglione, C., Gache, C., 1992a. Spatial and temporal expression pattern during sea urchin embryogenesis of a gene coding for a protease homologous to the human protein BMP-1 and to the product of the *Drosophila* dorsal–ventral patterning gene tolloid. Development 114, 147–163.

Lepage, T., Sardet, C., Gache, C., 1992b. Spatial expression of the hatching enzyme gene in the sea urchin embryo. Dev. Biol. 150, 23–32.

Logan, C.Y., Miller, J.R., Ferkowicz, M.J., McClay, D.R., 1999. Nuclear beta-catenin is required to specify vegetal cell fates in the sea urchin embryo. Development 126, 345–357.

Marques, G., Musacchio, M., Shimell, M.J., Wunnenberg-Stapleton, K., Cho, K.W., O'Connor, M.B., 1997. Production of a DPP activity gradient in the early *Drosophila* embryo through the opposing actions of the SOG and TLD proteins. Cell 91, 417–426.

Mohrlen, F., Hutter, H., Zwilling, R., 2003. The astacin protein family in *Caenorhabditis elegans*. Eur. J. Biochem. 270, 4909–4920.

Mohrlen, F., Maniura, M., Plickert, G., Frohme, M., Frank, U., 2006. Evolution of astacin-like metalloproteases in animals and their function in development. Evol. Dev. 8, 223–231.

Minokawa, T., Rast, J.P., Arenas-Mena, C., Franco, C.B., Davidson, E.H., 2004. Expression patterns of four different regulatory genes that function during sea urchin development. Gene Exp. Patterns 4, 449–456.

Mullins, M.C., 1998. Holy Tolloido: Tolloid cleaves SOG/Chordin to free DPP/BMPs. Trends Genet. 14, 127–129.

Nicholson, A.C., Malik, S.B., Logsdon Jr., J.M., Van Meir, E.G., 2005. Functional evolution of ADAMTS genes: evidence from analyses of phylogeny and gene organization. BMC Evol. Biol. 5, 11.

Pappano, W.N., Steiglitz, B.M., Scott, I.C., Keene, D.R., Greenspan, D.S., 2003. Use of Bmp1/Tll1 doubly homozygous null mice and proteomics to identify and validate in vivo substrates of bone morphogenetic protein 1/tolloid-like metalloproteinases. Mol. Cell. Biol. 23, 4428–4438.

Piccolo, S., Agius, E., Lu, B., Goodman, S., Dale, L., De Robertis, E.M., 1997. Cleavage of Chordin by Xolloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. Cell 91, 407–416.

Reynolds, S.D., Angerer, L.M., Palis, J., Nasir, A., Angerer, R.C., 1992. Early

mRNAs, spatially restricted along the animal–vegetal axis of sea urchin embryos, include one encoding a protein related to tolloid and BMP-1. Development 114, 769–786.

Rise, M., Burke, R.D., 2002. SpADAM, a sea urchin ADAM, has conserved structure and expression. Mech. Dev. 117, 275–281.

Sarras Jr., M.P., 1996. BMP-1 and the astacin family of metalloproteinases: a potential link between the extracellular matrix, growth factors and pattern formation. BioEssays 18, 439–442.

Stocker, W., Gomis-Ruth, F.X., Bode, W., Zwilling, R., 1993. Implications of the three-dimensional structure of astacin for the structure and function of the astacin family of zinc-endopeptidases. Eur. J. Biochem. 214, 215–231.

Samanta, M., Tongprasit, W., Istrrail, S., Cameron, R., Tu, Q., Davidson, E., Stole, V., in press . A high-resolution transcriptome map of the sea urchin embryo. Science.

Sherwood, D.R., McClay, D.R., 1999. LvNotch signaling mediates secondary mesenchyme specification in the sea urchin embryo. Development 126, 1703–1713.

Sherwood, D.R., McClay, D.R., 2001. LvNotch signaling plays a dual role in regulating the position of the ectoderm–endoderm boundary in the sea urchin embryo. Development 128, 2221–2232.

Sweet, H.C., Hodor, P.G., Ettensohn, C.A., 1999. The role of micromere signaling in Notch activation and mesoderm specification during sea urchin embryogenesis. Development 126, 5255–5265.

Takahara, K., Lyons, G.E., Greenspan, D.S., 1994. Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encoded by alternatively spliced transcripts which are differentially expressed in some tissues. J. Biol. Chem. 269, 32572–32578.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Trackman, P.C., 2005. Diverse biological functions of extracellular collagen processing enzymes. J. Cell. Biochem. 96, 927–937.

Vonica, A., Weng, W., Gumbiner, B.M., Venuti, J.M., 2000. TCF is the nuclear effector of the beta–catenin signal that patterns the sea urchin animal–vegetal axis. Dev. Biol. 217, 230–243.

Vu, T.H., Werb, Z., 2000. Matrix metalloproteinases: effectors of development and normal physiology. Genes Dev. 14, 2123–2133.

Wessel, G.M., McClay, D.R., 1987. Gastrulation in the sea urchin embryo requires the deposition of crosslinked collagen within the extracellular matrix. Dev. Biol. 121, 149–165.

Wikramanayake, A.H., Huang, L., Klein, W.H., 1998. β-Catenin is essential for patterning the maternally specified animal-vegetal axis in the sea urchin embryo. Proc. Natl. Acad. Sci. U.S.A. 95, 9343–9348.

Yiallouros, I., Kappelhoff, R., Schilling, O., Wegmann, F., Helms, M.W., Auge, A., Brachtendorf, G., Berkhoff, E.G., Beermann, B., Hinz, H.J., Konig, S., Peter-Katalinic, J., Stocker, W., 2002. Activation mechanism of pro-astacin: role of the pro-peptide, tryptic and autoproteolytic cleavage and importance of precise amino-terminal processing. J. Mol. Biol. 324, 237–246.

Zito, F., Costa, C., Sciarrino, S., Poma, V., Russo, R., Angerer, L.M., Matranga, V., 2003. Expression of univin, a TGF-beta growth factor, requires ectoderm-ECM interaction and promotes skeletal growth in the sea urchin embryo. Dev. Biol. 264, 217–227.

Zito, F., Nakano, E., Sciarrino, S., Matranga, V., 2000. Regulative specification of ectoderm in skeleton disrupted sea urchin embryos treated with monoclonal antibody to Pl-nectin. Dev. Growth Differ. 42, 499–506.